ADVANCED QUANTUM MECHANICS

Pieter Kok
University of Sheffield

University of Sheffield

Advanced Quantum Mechanics

Pieter Kok

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the hundreds of other texts available within this powerful platform, it is freely available for reading, printing and "consuming." Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects.

Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.

The LibreTexts mission is to unite students, faculty and scholars in a cooperative effort to develop an easy-to-use online platform for the construction, customization, and dissemination of OER content to reduce the burdens of unreasonable textbook costs to our students and society. The LibreTexts project is a multi-institutional collaborative venture to develop the next generation of open-access texts to improve postsecondary education at all levels of higher learning by developing an Open Access Resource environment. The project currently consists of 14 independently operating and interconnected libraries that are constantly being optimized by students, faculty, and outside experts to supplant conventional paper-based books. These free textbook alternatives are organized within a central environment that is both vertically (from advance to basic level) and horizontally (across different fields) integrated.

The LibreTexts libraries are Powered by MindTouch® and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739. Unless otherwise noted, LibreTexts content is licensed by CC BY-NC-SA 3.0.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org. More information on our activities can be found via Facebook (https://facebook.com/Libretexts), Twitter (https://twitter.com/libretexts), or our blog (http://Blog.Libretexts.org).

This text was compiled on 04/09/2023

TABLE OF CONTENTS

Licensing

1: Linear Vector Spaces and Hilbert Space

- 1.1: Linear Vector Spaces
- 1.2: Operators in Hilbert Space
- 1.3: Hermitian and Unitary Operators
- 1.4: Projection Operators and Tensor Products
- 1.5: The Trace and Determinant of an Operator

2: The Postulates of Quantum Mechanics

3: Schrödinger and Heisenberg Pictures

4: Mixed States and the Density Operator

- 4.1: Mixed States
- 4.2: Decoherence
- 4.3: Imperfect Measurements

5: Composite Systems and Entanglement

- 5.1: Composite Systems
- 5.2: Entanglement
- 5.3: Quantum teleportation

6: Evolution of Open Quantum Systems

- 6.1: The Lindblad Equation
- 6.2: Positive and Completely Positive Maps
- 6.3: Bra Vectors and the Inner Product
- 6.4: Normalization and Orthogonality
- 6.5: Interpreting the Inner Product
  - 6.5.1: Propagating Amplitudes

7: Orbital Angular Momentum and Spin

- 7.1: Orbital Angular Momentum
- 7.2: Spin
- 7.3: Total Angular Momentum
- 7.4: Composite Systems with Angular Momentum

8: Identical Particles

- 8.1: Symmetric and Anti-symmetric States
- 8.2: Creation and Annihilation Operators
- 8.3: Observables Based on Creation and Annihilation Operators
- 8.4: Bose-Einstein and Fermi-Dirac Statistics

9: Many-Body Problems in Quantum Mechanics

- 9.1: Interacting Electrons in Atomic Shells
- 9.2 Spin Waves in Solids
- 9.3: An Atom in a Cavity
- 9.4: Outlook: Quantum Field Theory

10: Pauli Spin Matrices

- 10.1: Spin Operators
- 10.2: Expectation Values
- 10.3: Total Angular Momentum

11: Noncommuting Operators and Uncertainty

- 11.1: Eigenstates and Commuting Operators
- 11.2: Non-Commuting Operators
- 11.3: Quantifying Uncertainty
  - 11.3.1: Mean and Variance
  - 11.3.2: Uncertainty in Quantum Mechanics
- 11.4: The Heisenberg Uncertainty Principle
  - 11.4.1: An Alternate Formulation
  - 11.4.2: Vacuum Energy

12: Multiple Particle States

- 12.1: Indistinguishable Particles
- 12.2: Notating Multiple Particle States
- 12.3: The Exchange Operator
- 12.4: Fermions and Bosons
- 12.5: The Pauli Exclusion Principle
- 12.6: Entangled Particles

13: The Schrödinger Equation

- 13.1: Where we are so far
- 13.2: Stating the Equation
- 13.3: Free Particles and the de Broglie Wavelength
- 13.4: Quantized Energy Levels in Bound Systems
- 13.5: The Simple Harmonic Oscillator
- 13.6: The Hydrogen Atom
- 13.7: Interpretation of the Wave Function ψ(x)

14: Atomic Orbitals

- 14.1: The Schrödinger Equation
- 14.2: The Orbitals
- 14.3: Visualizing Orbitals
  - 14.3.1: s Orbitals
  - 14.3.2: p Orbitals
  - 14.3.3: d Orbitals

15: The Periodic Table of the Elements

- 15.1: Interacting Electrons, Energy Levels, and Filled Shells
- 15.2: Filling Up Orbitals
- 15.3: Reading a Periodic Table
  - 15.3.1: Electronic Configuration

16: Matter

- 16.1: The Standard Model of Particle Physics
- 16.2: Nuclei and Atoms
- 16.3: Molecules
- 16.4: Solids
- 16.5: Liquids, Gasses, and Plasmas
  - 16.5.1: Quantum Gasses
- 16.6: Planets, Stars, Galaxies, and Clusters
- 16.7: Dark Matter and Dark Energy

Index

Glossary

Detailed Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

CHAPTER OVERVIEW

1: Linear Vector Spaces and Hilbert Space

The modern version of quantum mechanics was formulated in 1932 by John von Neumann in his famous book Mathematical Foundations of Quantum Mechanics, and it unifies Schrödingers wave theory with the matrix mechanics of Heisenberg, Born, and Jordan. The theory is framed in terms of linear vector spaces, so the first couple of lectures we have to remind ourselves of the relevant mathematics.

1.1: Linear Vector Spaces
1.2: Operators in Hilbert Space
1.3: Hermitian and Unitary Operators
1.4: Projection Operators and Tensor Products
1.5: The Trace and Determinant of an Operator

This page titled 1: Linear Vector Spaces and Hilbert Space is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

1.1: Linear Vector Spaces

Consider a set of vectors, denoted by $|\psi\rangle, |\phi\rangle$, etc., and the complex numbers a, b, c, etc. A linear vector space $\mathcal{V}$ is a mathematical structure of vectors and numbers that obeys the following rules:

1. $|\psi\rangle + |\phi\rangle = |\phi\rangle + |\psi\rangle$ (commutativity),
2. $|\psi\rangle + (|\phi\rangle + |\chi\rangle) = (|\psi\rangle + |\phi\rangle) + |\chi\rangle$ (associativity),
3. $a(|\psi\rangle + |\phi\rangle) = a|\psi\rangle + a|\phi\rangle$ (linearity),
4. $(a+b)|\psi\rangle = a|\psi\rangle + b|\psi\rangle$ (linearity),
5. $a(b|\phi\rangle) = (ab)|\phi\rangle$.

There is also a null vector 0 such that $|\psi\rangle + 0 = |\psi\rangle$, and for every $|\psi\rangle$ there is a vector $|\phi\rangle$ such that $|\psi\rangle + |\phi\rangle = 0$.

For each vector $|\phi\rangle$ there is a dual vector $\langle\phi|$, and the set of dual vectors also form a linear vector space $\mathcal{V}^*$. There is an inner product between vectors from $\mathcal{V}$ and $\mathcal{V}^*$ denoted by $\langle\psi \mid \phi\rangle$. The inner product has the following properties:

1. $\langle\psi \mid \phi\rangle = \langle\phi \mid \psi\rangle^*$,
2. $\langle\psi \mid \psi\rangle \geq 0$,
3. $\langle\psi \mid \psi\rangle = 0 \Leftrightarrow |\psi\rangle = 0$,
4. $|\psi\rangle = c_1 |\psi_1\rangle + c_2 |\psi_2\rangle \Rightarrow \langle\phi \mid \psi\rangle = c_1 \langle\phi \mid \psi_1\rangle + c_2 \langle\phi \mid \psi_2\rangle$,
5. $\|\phi\| \equiv \sqrt{\langle\phi \mid \phi\rangle}$ is the norm of $|\phi\rangle$.

If $\|\phi\| = 1$, the vector $|\phi\rangle$ is a unit vector. The set of unit vectors $\left\{e^{i\varphi}|\psi\rangle\right\}$ with $\varphi \in [0, 2\pi)$ form a so-called ray in the linear vector space. A linear vector space that has a norm $\|.\|$ (there are many different ways we can define a norm) is called a Hilbert space. We will always assume that the linear vector spaces are Hilbert spaces.

For linear vector spaces with an inner product we can derive the Cauchy-Schwarz inequality, also known as the Schwarz inequality:

$$|\langle\phi \mid \psi\rangle|^2 \leq \langle\psi \mid \psi\rangle\langle\phi \mid \phi\rangle \tag{1.1}$$

This is a very important relation, since it requires only the inner product structure. Relations that are based on this inequality, such as the Heisenberg uncertainty relation between observables, therefore have a very general validity.

If two vectors have an inner product equal to zero, then these vectors are called orthogonal. This is the definition of orthogonality. When these vectors are also unit vectors, they are called orthonormal. A set of vectors $|\phi_1\rangle, |\phi_2\rangle, \ldots |\phi_N\rangle$ are linearly independent if

$$\sum_j a_j |\phi_j\rangle = 0 \tag{1.2}$$

implies that all $a_j = 0$. The maximum number of linearly independent vectors in $\mathcal{V}$ is the dimension of $\mathcal{V}$. Orthonormal vectors form a complete orthonormal basis for $\mathcal{V}$ if any vector can be written as

$$|\psi\rangle = \sum_{k=1}^{N} c_k |\phi_k\rangle \tag{1.3}$$

and $\langle\phi_j \mid \phi_k\rangle = \delta_{jk}$. We can take the inner product of $|\psi\rangle$ with any of the basis vectors $|\phi_j\rangle$ to obtain

$$\langle\phi_j \mid \psi\rangle = \sum_{k=1}^{N} c_k \langle\phi_j \mid \phi_k\rangle = \sum_{k=1}^{N} c_k \delta_{jk} = c_j \tag{1.4}$$

Substitute this back into the expansion of $|\psi\rangle$, and we find

$$|\psi\rangle = \sum_{k=1}^{N} |\phi_k\rangle \langle\phi_k \mid \psi\rangle \tag{1.5}$$

Therefore $\sum_k |\phi_k\rangle \langle\phi_k|$ must act like the identity. In fact, this gives us an important clue that operators of states must take the general form of sums over objects like $|\phi\rangle\langle\chi|$.

This page titled 1.1: Linear Vector Spaces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

1.2: Operators in Hilbert Space

The objects $|\psi\rangle$ are vectors in a Hilbert space. We can imagine applying rotations of the vectors, rescaling, permutations of vectors in a basis, and so on. These are described mathematically as operators, and we denote them by capital letters A, B, C, etc. In general we write

$$A|\phi\rangle = |\psi\rangle \tag{1.6}$$

for some $|\phi\rangle, |\psi\rangle \in \mathcal{V}$. It is important to remember that operators act on all the vectors in Hilbert space. Let $\{|\phi_j\rangle\}_j$ be an orthonormal basis. We can calculate the inner product between the vectors $|\phi_j\rangle$ and A $|\phi_k\rangle$:

$$\langle\phi_j|\left(A\,|\phi_k\rangle\right) = \langle\phi_j|A|\phi_k\rangle \equiv A_{jk} \tag{1.7}$$

The two indices indicate that operators are matrices.

As an example, consider two vectors, written as two-dimensional column vectors

$$|\phi_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\phi_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.8}$$

and suppose that

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \tag{1.9}$$

We calculate

$$A_{11} = \langle\phi_1|A|\phi_1\rangle = (1,0)\cdot\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = (1,0)\cdot\begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \tag{1.10}$$

Similarly, we can calculate that $A_{22} = 3$, and $A_{12} = A_{21} = 0$ (check this). We therefore have that $A\,|\phi_1\rangle = 2\,|\phi_1\rangle$ and $A\,|\phi_2\rangle = 3\,|\phi_2\rangle$.

Complex numbers a have complex conjugates a^* and vectors $|\psi\rangle$ have dual vectors $\langle\phi|$. Is there an equivalent for operators? The answer is yes, and it is called the adjoint, or Hermitian conjugate, and is denoted by a dagger †. The natural way to define it is according to the rule

$$\langle\psi|A|\phi\rangle^* = \left\langle\phi\left|A^\dagger\right|\psi\right\rangle \tag{1.11}$$

for any $|\phi\rangle$ and $|\psi\rangle$. In matrix notation, and given an orthonormal basis $\{|\phi_j\rangle\}_j$, this becomes

$$\langle\phi_j|A|\phi_k\rangle^* = A^*_{jk} = \left\langle\phi_k\left|A^\dagger\right|\phi_j\right\rangle = A^\dagger_{kj} \tag{1.12}$$

So the matrix representation of the adjoint $A^\dagger$ is the transpose and the complex conjugate of the matrix A, as given by $\left(A^\dagger\right)_{jk} = A^*_{kj}$. The adjoint has the following properties:

1. $(cA)^\dagger = c^* A^\dagger$,
2. $(AB)^\dagger = B^\dagger A^\dagger$,
3. $(|\phi\rangle)^\dagger = \langle\phi|$.

Note the order of the operators in 2: AB is generally not the same as BA. The difference between the two is called the commutator, denoted by

$$[A, B] = AB - BA \tag{1.13}$$

For example, we can choose

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.14}$$

which leads to

$$\begin{aligned}[A,B] &= \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} - \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} - \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} \\ &= \begin{pmatrix}0 & -2\\ 2 & 0\end{pmatrix} \neq 0.\end{aligned} \tag{1.15}$$

Many, but not all, operators have an inverse. Let $A|\phi\rangle = |\psi\rangle$ and $B|\psi\rangle = |\phi\rangle$. Then we have

$$BA|\phi\rangle = |\phi\rangle \quad \text{and} \quad AB|\psi\rangle = |\psi\rangle \tag{1.16}$$

If Eq. (1.16) holds true for all $|\phi\rangle$ and $|\psi\rangle$, then B is the inverse of A, and we write $B = A^{-1}$. An operator that has an inverse is called invertible. Another important property that an operator may possess is positivity. An operator is positive if

$$\langle\phi|A|\phi\rangle \geq 0 \quad \text{for all } |\phi\rangle \tag{1.17}$$

We also write this as $A \geq 0$.

From the matrix representation of operators you can easily see that the operators themselves form a linear vector space:

1. $A+B = B+A$,
2. $A+(B+C) = (A+B)+C$,
3. $a(A+B) = aA+aB$,
4. $(a+b)A = aA+bA$,
5. $a(bA) = (ab)A$.

We can also define the operator norm $\|A\|$ according to

$$\|A\| = \sqrt{\mathrm{Tr}\left(A^\dagger A\right)} \equiv \sqrt{\sum_{ij} A^*_{ij}A_{ji}}, \tag{1.18}$$

which means that the linear vector space of operators is again a Hilbert space. The symbol Tr(.) denotes the trace of an operator, and we will return to this special operator property later in this section.

Every operator has a set of vectors for which

$$A\,|a_j\rangle = a_j\,|a_j\rangle\,, \qquad \text{with } a_j \in \mathbb{C} \tag{1.19}$$

This is called the eigenvalue equation (or eigenequation) for A, and the vectors $|a_j\rangle$ are the eigenvectors. The complex numbers a_j are eigenvalues. In the basis of eigenvectors, the matrix representation of A becomes

$$A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_N \end{pmatrix} \tag{1.20}$$

When some of the a_js are the same, we speak of degenerate eigenvalues. When there are n identical eigenvalues, we have n-fold degeneracy. The eigenvectors corresponding to this eigenvalue then span an n-dimensional subspace of the vector space. We will return to subspaces shortly, when we introduce projection operators.

For any orthonormal basis $\{|\phi_j\rangle\}_j$ we have

$$\langle\phi_j|A|\phi_k\rangle = A_{jk} \tag{1.21}$$

which can be written in the form

$$A = \sum_{jk} A_{jk}\,|\phi_j\rangle\,\langle\phi_k| \tag{1.22}$$

For the special case where $|\phi_j\rangle = |a_j\rangle$ this reduces to

$$A = \sum_{j} a_j\,|a_j\rangle\,\langle a_j| \tag{1.23}$$

This is the spectral decomposition of A. When all a_j are equal, we have complete degeneracy over the full vector space, and the operator becomes proportional (up to a factor a_j) to the identity $\mathbb{I}$.

Note that this is independent of the basis $\{|a_j\rangle\}$. As a consequence, for any orthonormal basis $\{|\phi_j\rangle\}$ we have

$$\mathbb{I} = \sum_j |\phi_j\rangle \langle\phi_j| \tag{1.24}$$

This is the completeness relation, and we will use this many times in our calculations.

Lemma

If two non-degenerate operators commute $([A, B] = 0)$, then they have a common set of eigenvectors.

Proof

Let $A = \sum_k a_k |a_k\rangle \langle a_k|$ and $B = \sum_{jk} B_{jk} |a_j\rangle \langle a_k|$. We can choose this without loss of generality: we write both operators in the eigenbasis of A. Furthermore, $[A, B] = 0$ implies that $AB = BA$.

$$\begin{aligned} AB &= \sum_{klm} a_k B_{lm} |a_k\rangle\langle a_k \mid a_l\rangle\langle a_m| = \sum_{lm} a_l B_{lm} |a_l\rangle\langle a_m| \\ BA &= \sum_{klm} a_k B_{lm} |a_l\rangle\langle a_m \mid a_k\rangle\langle a_k| = \sum_{lm} a_m B_{lm} |a_l\rangle \langle a_m| \end{aligned} \tag{1.25}$$

Therefore

$$[A, B] = \sum_{lm} (a_l - a_m) B_{lm} |a_l\rangle \langle a_m| = 0 \tag{1.26}$$

If $a_l \neq a_m$ for $l \neq m$, then $B_{lm} = 0$, and $B_{lm} \propto \delta_{lm}$. Therefore $\{|a_j\rangle\}$ is an eigenbasis for B.

The proof of the converse is left as an exercise. It turns out that this is also true when A and/or B are degenerate.

This page titled 1.2: Operators in Hilbert Space is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

1.3: Hermitian and Unitary Operators

Next, we will consider two special types of operators, namely Hermitian and unitary operators. An operator A is Hermitian if and only if $A^\dagger = A$.

Lemma

An operator is Hermitian if and only if it has real eigenvalues: $A^\dagger = A \Leftrightarrow a_j \in \mathbb{R}$.

Proof

The eigenvalue equation of A implies that

$$A\left|a_j\right\rangle = a_j\left|a_j\right\rangle \Rightarrow \left\langle a_j\right| A^\dagger = a_j^*\left\langle a_j\right|, \tag{1.27}$$

which means that $\langle a_j|A|a_j\rangle = a_j$ and $\left\langle a_j\left|A^\dagger\right|a_j\right\rangle = a_j^*$. It is now straightforward to show that $A = A^\dagger$ implies $a_j = a_j^*$, or $a_j \in \mathbb{R}$. Conversely, $a_j \in \mathbb{R}$ implies $a_j = a_j^*$, and

$$\langle a_j|A|a_j\rangle = \left\langle a_j\left|A^\dagger\right|a_j\right\rangle \tag{1.28}$$

Let $|\psi\rangle = \sum_k c_k\left|a_k\right\rangle$. Then

$$\begin{aligned}\langle\psi|A|\psi\rangle &= \sum_j |c_j|^2\langle a_j|A|a_j\rangle = \sum_j |c_j|^2\left\langle a_j\left|A^\dagger\right|a_j\right\rangle = \sum_j |c_j|^2\left\langle a_j\left|A^\dagger\right|a_j\right\rangle \\ &= \left\langle\psi\left|A^\dagger\right|\psi\right\rangle\end{aligned} \tag{1.29}$$

for all $|\psi\rangle$, and therefore $A = A^\dagger$.

Next, we construct the exponent of an operator A according to $U = \exp(icA)$. We have included the complex number c for completeness. At first sight, you may wonder what it means to take the exponent of an operator. Recall, however, that the exponent has a power expansion:

$$U = \exp(icA) = \sum_{n=0}^{\infty}\frac{(ic)^n}{n!}A^n \tag{1.30}$$

The n^{th} power of an operator is straightforward: just multiply A n times with itself. The expression in Eq. (1.30) is then well defined, and the exponent is taken as an abbreviation of the power expansion. In general, we can construct any function of operators, as long as we can define the function in terms of a power expansion:

$$f(A) = \sum_{n=0}^{\infty} f_n A^n \tag{1.31}$$

This can also be extended to functions of multiple operators, but now we have to be very careful in the case where these operators do not commute. Familiar rules for combining normal functions no longer apply (see exercise 4b).

We can construct the adjoint of the operator U according to

$$U^\dagger = \left(\sum_{n=0}^{\infty}\frac{(ic)^n}{n!}A^n\right)^\dagger = \sum_{n=0}^{\infty}\frac{(-ic^*)^n}{n!}A^{\dagger n} = \exp\left(-ic^*A^\dagger\right) \tag{1.32}$$

In the special case where $A = A^\dagger$ and c is real, we calculate

$$UU^\dagger = \exp(icA)\exp\left(-ic^*A^\dagger\right) = \exp(icA)\exp(-icA) = \exp[ic(A-A)] = \mathbb{I}, \tag{1.33}$$

since A commutes with itself. Similarly, $U^\dagger U = \mathbb{I}$. Therefore, $U^\dagger = U^{-1}$, and an operator with this property is called unitary. Each unitary operator can be generated by a Hermitian (self-adjoint) operator A and a real number c. A is called the generator of U. We often write $U = U_A(c)$. Unitary operators are basis transformations.

This page titled 1.3: Hermitian and Unitary Operators is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

1.4: Projection Operators and Tensor Products

We can combine two linear vector spaces $\mathscr{U}$ and $\mathscr{V}$ into a new linear vector space $\mathscr{W} = \mathscr{U} \oplus \mathscr{V}$. The symbol $\oplus$ is called the direct sum. The dimension of $\mathscr{W}$ is the sum of the dimensions of $\mathscr{U}$ and $\mathscr{V}$:

$$\dim \mathscr{W} = \dim \mathscr{U} + \dim \mathscr{V} \tag{1.34}$$

A vector in $\mathscr{W}$ can be written as

$$|\Psi\rangle_{\mathscr{W}} = |\psi\rangle_{\mathscr{U}} + |\phi\rangle_{\mathscr{V}}, \tag{1.35}$$

where $|\psi\rangle_{\mathscr{U}}$ and $|\phi\rangle_{\mathscr{V}}$ are typically not normalized (i.e., they are not unit vectors). The spaces $\mathscr{U}$ and $\mathscr{V}$ are so-called subspaces of $\mathscr{W}$.

As an example, consider the three-dimensional Euclidean space spanned by the Cartesian axes x, y, and z. The xy-plane is a two-dimensional subspace of the full space, and the z-axis is a one-dimensional subspace. Any three-dimensional form can be projected onto the xy-plane by setting the z component to zero. Similarly, we can project onto the z-axis by setting the x and y coordinates to zero. A projector is therefore associated with a subspace. It acts on a vector in the full space, and forces all components to zero, except those of the subspace it projects onto.

The formal definition of a projector $P_{\mathscr{U}}$ on $\mathscr{U}$ is given by

$$P_{\mathscr{U}} |\Psi\rangle_{\mathscr{W}} = |\psi\rangle_{\mathscr{U}} \tag{1.36}$$

This is equivalent to requiring that $P_{\mathscr{U}}^2 = P_{\mathscr{U}}$, $P_{\mathscr{U}}^2 = P_{\mathscr{U}}$, or $\boldsymbol{P}_{\mathscr{U}}$ is idempotent. One-dimensional projectors can be written as

$$P_j = |\phi_j\rangle \langle\phi_j| \tag{1.37}$$

Two projectors P_1 and P_2 are orthogonal is $P_1 P_2 = 0$. If $P_1 P_2 = 0$, then $P_1 + P_2$ is another projector:

$$(P_1 + P_2)^2 = P_1^2 + P_1 P_2 + P_2 P_1 + P_2^2 = P_1^2 + P_2^2 = P_1 + P_2 \tag{1.38}$$

When P_1 and P_2 commute but are non-orthogonal (i.e., they overlap), the general projector onto their combined subspace is

$$P_{1+2} = P_1 + P_2 - P_1 P_2 \tag{1.39}$$

(Prove this.) The orthocomplement of P is $\mathbb{I} - P$, which is also a projector:

$$P(\mathbb{I} - P) = P - P^2 = P - P = 0 \quad \text{and} \quad (\mathbb{I} - P)^2 = \mathbb{I} - 2P + P^2 = \mathbb{I} - P \tag{1.40}$$

Another way to combine two vector spaces $\mathscr{U}$ and $\mathscr{V}$ is via the tensor product: $\mathscr{W} = \mathscr{U} \otimes \mathscr{V}$, where the symbol $\otimes$ is called the direct product or tensor product. The dimension of the space $\mathscr{W}$ is then

$$\dim \mathscr{W} = \dim \mathscr{U} \cdot \dim \mathscr{V} \tag{1.41}$$

Let $|\psi\rangle \in \mathscr{U}$ and $|\phi\rangle \in \mathscr{V}$. Then

$$|\psi\rangle \otimes |\phi\rangle \in \mathscr{W} = \mathscr{U} \otimes \mathscr{V} \tag{1.42}$$

If $|\psi\rangle = \sum_j a_j |\psi_j\rangle$ and $|\phi\rangle = \sum_j b_j |\phi_j\rangle$, then the tensor product of these vectors can be written as

$$|\psi\rangle \otimes |\phi\rangle = \sum_{jk} a_j b_k |\psi_j\rangle \otimes |\phi_k\rangle = \sum_{jk} a_j b_k |\psi_j\rangle |\phi_k\rangle = \sum_{jk} a_j b_k |\psi_j, \phi_k\rangle, \tag{1.43}$$

where we introduced convenient abbreviations for the tensor product notation. The inner product of two vectors that are tensor products is

$$(\langle\psi_1| \otimes \langle\phi_1|)(|\psi_2\rangle \otimes |\phi_2\rangle) = \langle\psi_1 \mid \psi_2\rangle \langle\phi_1 \mid \phi_2\rangle \tag{1.44}$$

Operators also obey the tensor product structure, with

$$(A \otimes B)|\psi\rangle \otimes |\phi\rangle = (A|\psi\rangle) \otimes (B|\phi\rangle) \tag{1.45}$$

and

$$(A \otimes B)(C \otimes D)|\psi\rangle \otimes |\phi\rangle = (AC|\psi\rangle) \otimes (BD|\phi\rangle) \tag{1.46}$$

General rules for tensor products of operators are

1. $A \otimes 0 = 0$ and $0 \otimes B = 0$,
2. $\mathbb{I} \otimes \mathbb{I} = \mathbb{I}$,
3. $(A_1 + A_2) \otimes B = A_1 \otimes B + A_2 \otimes B$,
4. $aA \otimes bB = (ab) A \otimes B$,
5. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$,
6. $(A \otimes B)^{\dagger} = A^{\dagger} \otimes B^{\dagger}$.

Note that the last rule preserves the order of the operators. In other words, operators always act on their own space. Often, it is understood implicitly which operator acts on which subspace, and we will write $A \otimes \mathbb{I} = A$ and $\mathbb{I} \otimes B = B$. Alternatively, we can add subscripts to the operator, e.g., $A_{\mathscr{U}}$ and $B_{\mathscr{V}}$.

As a practical example, consider two two-dimensional operators

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \tag{1.47}$$

with respect to some orthonormal bases $\{|a_1\rangle, |a_2\rangle\}$ and $\{|b_1\rangle, |b_2\rangle\}$ for A and B, respectively (not necessarily eigenbases). The question is now: what is the matrix representation of $A \otimes B$? Since the dimension of the new vector space is the product of the dimensions of the two vector spaces, we have $\dim \mathscr{W} = 2 \cdot 2 = 4$. A natural basis for $A \otimes B$ is then given by $\{|a_j, b_k\rangle\}_{jk}$, with j, $k = 1, 2$, or

$$|a_1\rangle |b_1\rangle, \quad |a_1\rangle |b_2\rangle, \quad |a_2\rangle |b_1\rangle, \quad |a_2\rangle |b_2\rangle \tag{1.48}$$

We can construct the matrix representation of $A \otimes B$ by applying this operator to the basis vectors in Eq. (1.48), using

$$A |a_j\rangle = A_{1j} |a_1\rangle + A_{2j} |a_2\rangle \quad \text{and} \quad B |a_k\rangle = B_{1k} |b_1\rangle + B_{2k} |b_2\rangle \tag{1.49}$$

which leads to

$$\begin{aligned} A \otimes B |a_1, b_1\rangle &= (A_{11} |a_1\rangle + A_{21} |a_2\rangle)(B_{11} |b_1\rangle + B_{21} |b_2\rangle) \\ A \otimes B |a_1, b_2\rangle &= (A_{11} |a_1\rangle + A_{21} |a_2\rangle)(B_{12} |b_1\rangle + B_{22} |b_2\rangle) \\ A \otimes B |a_2, b_1\rangle &= (A_{12} |a_1\rangle + A_{22} |a_2\rangle)(B_{11} |b_1\rangle + B_{21} |b_2\rangle) \\ A \otimes B |a_2, b_2\rangle &= (A_{12} |a_1\rangle + A_{22} |a_2\rangle)(B_{12} |b_1\rangle + B_{22} |b_2\rangle) \end{aligned} \tag{1.50}$$

Looking at the first line of Eq. (1.50), the basis vector $|a_1, b_1\rangle$ gets mapped to all basis vectors:

$$A \otimes B |a_1, b_1\rangle = A_{11} B_{11} |a_1, b_1\rangle + A_{11} B_{21} |a_1, b_2\rangle + A_{21} B_{11} |a_2, b_1\rangle + A_{21} B_{21} |a_2, b_2\rangle \tag{1.51}$$

Combining this into matrix form leads to

$$A \otimes B = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} & A_{12}B_{11} & A_{12}B_{12} \\ A_{11}B_{21} & A_{11}B_{22} & A_{12}B_{21} & A_{12}B_{22} \\ A_{21}B_{11} & A_{21}B_{12} & A_{22}B_{11} & A_{22}B_{12} \\ A_{21}B_{21} & A_{21}B_{22} & A_{22}B_{21} & A_{22}B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{pmatrix} \tag{1.52}$$

Recall that this is dependent on the basis that we have chosen. In particular, $A \otimes B$ may be diagonalized in some other basis.

This page titled 1.4: Projection Operators and Tensor Products is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

1.5: The Trace and Determinant of an Operator

There are two special functions of operators that play a key role in the theory of linear vector spaces. They are the trace and the determinant of an operator, denoted by $\mathrm{Tr}(A)$ and $\det(A)$, respectively. While the trace and determinant are most conveniently evaluated in matrix representation, they are independent of the chosen basis.

When we defined the norm of an operator, we introduced the trace. It is evaluated by adding the diagonal elements of the matrix representation of the operator:

$$\mathrm{Tr}(A) = \sum_j \langle \phi_j | A | \phi_j \rangle \,, \tag{1.53}$$

where $\{|\phi_j\rangle\}_j$ is any orthonormal basis. This independence means that the trace is an invariant property of the operator. Moreover, the trace has the following important properties:

1. If $A = A^\dagger$, then $\mathrm{Tr}(A)$ is real,
2. $\mathrm{Tr}(aA) = a\,\mathrm{Tr}(A)$,
3. $\mathrm{Tr}(A+B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$,
4. $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ (the "cyclic property").

The first property follows immediately when we evaluate the trace in the diagonal basis, where it becomes a sum over real eigenvalues. The second and third properties convey the linearity of the trace. The fourth property is extremely useful, and can be shown as follows:

$$\begin{aligned} \mathrm{Tr}(AB) &= \sum_j \langle \phi_j | AB | \phi_j \rangle = \sum_{jk} \langle \phi_j | A | \psi_k \rangle \langle \psi_k | B | \phi_j \rangle \\ &= \sum_{jk} \langle \psi_k | B | \phi_j \rangle \langle \phi_j | A | \psi_k \rangle = \sum_k \langle \psi_k | BA | \psi_k \rangle \\ &= \mathrm{Tr}(BA) \end{aligned} \tag{1.54}$$

This derivation also demonstrates the usefulness of inserting a resolution of the identity in strategic places. In the cyclic property, the operators A and B may be products of two operators, which then leads to

$$\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) = \mathrm{Tr}(CAB) \tag{1.55}$$

Any cyclic (even) permutation of operators under a trace gives rise to the same value of the trace as the original operator ordering.

Finally, we construct the partial trace of an operator that lives on a tensor product space. Suppose that $A \otimes B$ is an operator in the Hilbert space $\mathscr{H}_1 \otimes \mathscr{H}_2$. We can trace out Hilbert space $\mathscr{H}_1$, denoted by $\mathrm{Tr}_1(\,.\,)$:

$$\mathrm{Tr}_1(A \otimes B) = \mathrm{Tr}(A) B, \quad \text{or equivalently} \quad \mathrm{Tr}_1(A_1 B_2) = \mathrm{Tr}(A_1) B_2 \tag{1.56}$$

Taking the partial trace has the effect of removing the entire Hilbert space $\mathscr{H}_1$ from the description. It reduces the total vector space. The partial trace always carries an index, which determines which space is traced over.

The determinant of a 2×2 matrix is given by

$$\det(A) = \det \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = A_{11}A_{22} - A_{12}A_{21} \tag{1.57}$$

The determinant of higher-dimensional matrices can be defined recursively as follows: The topleft element of an $n \times n$ matrix defines an $(n-1) \times (n-1)$ matrix by removing the top row and the left column. Similarly, any other element in the left column defines an $(n-1) \times (n-1)$ matrix by removing the left column and the row of the element we chose. The determinant of the $n \times n$ matrix is then given by the top-left element times the determinant of the remaining $(n-1) \times (n-1)$ matrix, minus the product of the second element down in the left column and the remaining $(n-1) \times (n-1)$ matrix, plus the third element times the remaining matrix, etc.

The determinant of the product of matrices is equal to the product of the determinants of the matrices:

$$\det(AB) = \det(A)\det(B) \tag{1.58}$$

Moreover, if A is an invertible matrix, then we have

$$\det\left(A^{-1}\right) = \det(A)^{-1} \tag{1.59}$$

This leads to an important relation between similar matrices $A = X^{-1}BX$:

$$\begin{aligned} \det(A) &= \det\left(X^{-1}BX\right) = \det\left(X^{-1}\right)\det(B)\det(X) \\ &= \det(X)^{-1}\det(B)\det(X) = \det(B) \end{aligned} \tag{1.60}$$

In particular, this means that the determinant is independent of the basis in which the matrix is written, which means that it is an intrinsic property of the operator associated with that matrix.

Finally, here's a fun relation between the trace and the determinant of an operator:

$$\det[\exp(A)] = \exp[\mathrm{Tr}(A)] \tag{1.61}$$

? Exercises

1. Vectors and matrices:
   a. Are the following three vectors linearly dependent or independent: $a = (2, 3, -1), b = (0, 1, 2)$, and $c = (0, 0, -5)$?
   b. Consider the vectors $|\psi\rangle = 3i\,|\phi_1\rangle - 7i\,|\phi_2\rangle$ and $|\chi\rangle = |\phi_1\rangle + 2\,|\phi_2\rangle$, with $\{|\phi_i\rangle\}$ an orthonormal basis. Calculate the inner product between $|\psi\rangle$ and $|\chi\rangle$, and show that they satisfy the Cauchy-Schwarz inequality.
   c. Consider the two matrices

$$A = \begin{pmatrix} 0 & i & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix} \tag{1.62}$$

   Calculate A^{-1} and BA^{-1}. Are they equal?
   d. Calculate $A \otimes B$ and $B \otimes A$, where $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
2. Operators:
   a. Which of these operators are Hermitian: $A + A^\dagger, i\left(A + A^\dagger\right), i\left(A - A^\dagger\right)$, and $A^\dagger A$?
   b. Prove that a shared eigenbasis for two operators A and B implies that $[A, B] = 0$.
   c. Let U be a transformation matrix that maps one complete orthonormal basis to another. Show that U is unitary
   d. How many real parameters completely determine a $d \times d$ unitary matrix?
3. Properties of the trace and the determinant:
   a. Calculate the trace and the determinant of the matrices A and B in exercise 1c.
   b. Show that the expectation value of A can be written as $\mathrm{Tr}(|\psi\rangle\langle\psi|A)$.
   c. Prove that the trace is independent of the basis.
4. Commutator identities.
   a. Let $F(t) = e^{At}e^{Bt}$. Calculate dF/dt and use $\left[e^{At}, B\right] = \left(e^{At}Be^{-At} - B\right)e^{At}$ to simplify your result.
   b. Let $G(t) = e^{At+Bt+f(t)H}$. Show by calculating dG/dt, and setting $dF/dt = dG/dt$ at $t = 1$, that the following operator identity

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]}, \tag{1.63}$$

   holds if A and B both commute with $[A, B]$. Hint: use the Hadamard lemma

$$e^{At}Be^{-At} = B + \frac{t}{1!}[A, B] + \frac{t^2}{2!}[A, [A, B]] + \ldots \tag{1.64}$$

   c. Show that the commutator of two Hermitian operators is anti-Hermitian $\left(A^\dagger = -A\right)$.
   d. Prove the commutator analog of the Jacobi identity

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \tag{1.65}$$

This page titled 1.5: The Trace and Determinant of an Operator is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

2: The Postulates of Quantum Mechanics

The entire structure of quantum mechanics (including its relativistic extension) can be formulated in terms of states and operations in Hilbert space. We need rules that map the physical quantities such as states, observables, and measurements to the mathematical structure of vector spaces, vectors and operators. There are several ways in which this can be done, and here we summarize these rules in terms of five postulates.

Postulate 1

A physical system is described by a Hilbert space $\mathscr{H}$, and the state of the system is represented by a ray with norm 1 in $\mathscr{H}$.

There are a number of important aspects to this postulate. First, the fact that states are rays, rather than vectors means that an overall phase $e^{i\varphi}$ of the state does not have any physically observable consequences, and $e^{i\varphi}|\psi\rangle$ represents the same state as $|\psi\rangle$. Second, the state contains all information about the system. In particular, there are no hidden variables in this standard formulation of quantum mechanics. Finally, the dimension of $\mathscr{H}$ may be infinite, which is the case, for example, when $\mathscr{H}$ is the space of square-integrable functions.

As an example of this postulate, consider a two-level quantum system (a qubit). This system can be described by two orthonormal states $|0\rangle$ and $|1\rangle$. Due to linearity of Hilbert space, the superposition $\alpha|0\rangle+\beta|1\rangle$ is again a state of the system if it has norm 1, or

$$(\alpha^*\langle 0|+\beta^*\langle 1|)(\alpha|0\rangle+\beta|1\rangle)=1 \quad \text{or} \quad |\alpha|^2+|\beta|^2=1 \tag{2.1}$$

This is called the superposition principle: any normalised superposition of valid quantum states is again a valid quantum state. It is a direct consequence of the linearity of the vector space, and as we shall see later, this principle has some bizarre consequences that have been corroborated in many experiments.

Postulate 2

Every physical observable A corresponds to a self-adjoint (Hermitian[1]) operator $\hat{A}$ whose eigenvectors form a complete basis.

We use a hat to distinguish between the observable and the operator, but usually this distinction is not necessary. In these notes, we will use hats only when there is a danger of confusion.

As an example, take the operator X:

$$X|0\rangle=|1\rangle \quad \text{and} \quad X|1\rangle=|0\rangle. \tag{2.2}$$

This operator can be interpreted as a bit flip of a qubit. In matrix notation the state vectors can be written as

$$|0\rangle=\begin{pmatrix}1\\0\end{pmatrix} \quad \text{and} \quad |1\rangle=\begin{pmatrix}0\\1\end{pmatrix}, \tag{2.3}$$

which means that X is written as

$$X=\begin{pmatrix}0&1\\1&0\end{pmatrix} \tag{2.4}$$

with eigenvalues ±1. The eigenstates of X are

$$|\pm\rangle=\frac{|0\rangle\pm|1\rangle}{\sqrt{2}}. \tag{2.5}$$

These states form an orthonormal basis.

Postulate 3

The eigenvalues of A are the possible measurement outcomes, and the probability of finding the outcome a_j in a measurement is given by the Born rule:

$$p\left(a_j\right)=\left|\langle a_j\mid\psi\rangle\right|^2, \tag{2.6}$$

where $|\psi\rangle$ is the state of the system, and $|a_j\rangle$ is the eigenvector associated with the eigenvalue a_j via $A|a_j\rangle = a_j|a_j\rangle$. If a_j is m-fold degenerate, then

$$p(a_j) = \sum_{l=1}^{m} |\langle a_j^{(l)} \mid \psi\rangle|^2, \tag{2.7}$$

where the $\left|a_j^{(l)}\right\rangle$ span the m-fold degenerate subspace

The expectation value of A with respect to the state of the system $|\psi\rangle$ is denoted by $\langle A\rangle$, and evaluated as

$$\langle A\rangle = \langle\psi|A|\psi\rangle = \langle\psi|(\sum_j a_j|a_j\rangle\langle a_j|)|\psi\rangle = \sum_j p(a_j)a_j \tag{2.8}$$

This is the weighted average of the measurement outcomes. The spread of the measurement outcomes (or the uncertainty) is given by the variance

$$(\Delta A)^2 = \left\langle (A - \langle A\rangle)^2 \right\rangle = \left\langle A^2 \right\rangle - \langle A\rangle^2 \tag{2.9}$$

So far we mainly dealt with discrete systems on finite-dimensional Hilbert spaces. But what about continuous systems, such as a particle in a box, or a harmonic oscillator? We can still write the spectral decomposition of an operator A but the sum must be replace by an integral:

$$A = \int da f_A(a)|a\rangle\langle a| \tag{2.10}$$

where $|a\rangle$ is an eigenstate of A. Typically, there are problems with the normalization of $|a\rangle$, which is related to the impossibility of preparing a system in exactly the state $|a\rangle$. We will not explore these subtleties further in this course, but you should be aware that they exist. The expectation value of A is

$$\langle A\rangle = \langle\psi|A|\psi\rangle = \int da f_A(a)\langle\psi \mid a\rangle\langle a \mid \psi\rangle \equiv \int da f_A(a)|\psi(a)|^2, \tag{2.11}$$

where we defined the wave function $\psi(a) = \langle a \mid \psi\rangle$, and $|\psi(a)|^2$ is properly interpreted as the probability density that you remember from second-year quantum mechanics.

The probability of finding the eigenvalue of an operator A in the interval a and $a + da$ given the state $|\psi\rangle$ is

$$\langle\psi|(|a\rangle\langle a|da)|\psi\rangle \equiv dp(a), \tag{2.12}$$

since both sides must be infinitesimal. We therefore find that

$$\frac{dp(a)}{da} = |\psi(a)|^2 \tag{2.13}$$

Postulate 4

The dynamics of quantum systems is governed by unitary transformations

We can write the state of a system at time t as $|\psi(t)\rangle$, and at some time $t_0 < t$ as $|\psi(t_0)\rangle$. The fourth postulate tells us that there is a unitary operator $U(t, t_0)$ that transforms the state at time t_0 to the state at time t:

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle \tag{2.14}$$

Since the evolution from time t to t is denoted by $U(t, t)$ and must be equal to the identity, we deduce that U depends only on time differences: $U(t, t_0) = U(t - t_0)$, and $U(0) = \mathbb{I}$.

As an example, let $U(t)$ be generated by a Hermitian operator A according to

$$U(t) = \exp\left(-\tfrac{i}{\hbar}At\right) \tag{2.15}$$

The argument of the exponential must be dimensionless, so A must be proportional to $\hbar$ times an angular frequency (in other words, an energy). Suppose that $|\psi(t)\rangle$ is the state of a qubit, and that $A = \hbar\omega X$. If $|\psi(0)\rangle = |0\rangle$ we want to calculate the state of

the system at time t. We can write

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle = \exp(-i\omega t X)|0\rangle = \sum_{n=0}^{\infty} \frac{(-i\omega t)^n}{n!} X^n \tag{2.16}$$

Observe that $X^2 = \mathbb{I}$, so we can separate the power series into even and odd values of n:

$$|\psi(t)\rangle = \sum_{n=0}^{\infty} \frac{(-i\omega t)^{2n}}{(2n)!}|0\rangle + \sum_{n=0}^{\infty} \frac{(-i\omega t)^{2n+1}}{(2n+1)!} X|0\rangle = \cos(\omega t)|0\rangle - i\sin(\omega t)|1\rangle \tag{2.17}$$

In other words, the state oscillates between $|0\rangle$ and $|1\rangle$.

The fourth postulate also leads to the Schrödinger equation. Let's take the infinitesimal form of Eq. (2.14):

$$|\psi(t+dt)\rangle = U(dt)|\psi(t)\rangle \tag{2.18}$$

We require that $U(dt)$ is generated by some Hermitian operator H:

$$U(dt) = \exp\left(-\frac{i}{\hbar} H dt\right) \tag{2.19}$$

H must have the dimensions of energy, so we identify it with the energy operator, or the Hamiltonian. We can now take a Taylor expansion of $|\psi(t+dt)\rangle$ to first order in dt:

$$|\psi(t+dt)\rangle = |\psi(t)\rangle + dt\frac{d}{dt}|\psi(t)\rangle + \ldots, \tag{2.20}$$

and we expand the unitary operator to first order in dt as well:

$$U(dt) = 1 - \frac{i}{\hbar} H dt + \ldots \tag{2.21}$$

We combine this into

$$|\psi(t)\rangle + dt\frac{d}{dt}|\psi(t)\rangle = \left(1 - \frac{i}{\hbar} H dt\right)|\psi(t)\rangle, \tag{2.22}$$

which can be recast into the Schrödinger equation:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle \tag{2.23}$$

Therefore, the Schrödinger equation follows directly from the postulates!

Postulate 5

If a measurement of an observable A yields an eigenvalue a_j, then immediately after the measurement, the system is in the eigenstate $|a_j\rangle$ corresponding to the eigenvalue

This is the infamous projection postulate, so named because a measurement "projects" the system to the eigenstate corresponding to the measured value. This postulate has as observable consequence that a second measurement immediately after the first will also find the outcome a_j. Each measurement outcome a_j corresponds to a projection operator P_j on the subspace spanned by the eigenvector(s) belonging to a_j. A (perfect) measurement can be described by applying a projector to the state, and renormalize:

$$|\psi\rangle \to \frac{P_j|\psi\rangle}{\|P_j|\psi\rangle\|} \tag{2.24}$$

This also works for degenerate eigenvalues.

We have established earlier that the expectation value of A can be written as a trace:

$$\langle A\rangle = \mathrm{Tr}(|\psi\rangle\langle\psi|A) \tag{2.25}$$

Now instead of the full operator A, we calculate the trace of $P_j = |a_j\rangle\langle a_j|$:

$$\langle P_j \rangle = \mathrm{Tr}(|\psi\rangle\langle\psi|P_j) = \mathrm{Tr}(|\psi\rangle \langle\psi \mid a_j\rangle \langle a_j|) = |\langle a_j \mid \psi\rangle|^2 = p\,(a_j) \tag{2.26}$$

So we can calculate the probability of a measurement outcome by taking the expectation value of the projection operator that corresponds to the eigenstate of the measurement outcome. This is one of the basic calculations in quantum mechanics that you should be able to do.

The Measurement Problem

The projection postulate is somewhat problematic for the interpretation of quantum mechanics, because it leads to the so-called **measurement problem**: Why does a measurement induce a non-unitary evolution of the system? After all, the measurement apparatus can also be described quantum mechanically[2] and then the system plus the measurement apparatus evolves unitarily. But then we must invoke a new device that measures the combined system and measurement apparatus. However, this in turn can be described quantum mechanically, and so on.

On the other hand, we do see definite measurement outcomes when we do experiments, so at some level the projection postulate is necessary, and somewhere there must be a "collapse of the wave function". Schrödinger already struggled with this question, and came up with his famous thought experiment about a cat in a box with a poison-filled vial attached to a Geiger counter monitoring a radioactive atom (Figure 1). When the atom decays, it will trigger the Geiger counter, which in turn causes the release of the poison killing the cat. When we do not look inside the box (more precisely: when no information about the atom-counter-vial-cat system escapes from the box), the entire system is in a quantum superposition. However, when we open the box, we do find the cat either dead or alive. One solution of the problem seems to be that the quantum state represents our knowledge of the system, and that looking inside the box merely updates our information about the atom, counter, vial and the cat. So nothing "collapses" except our own state of mind.

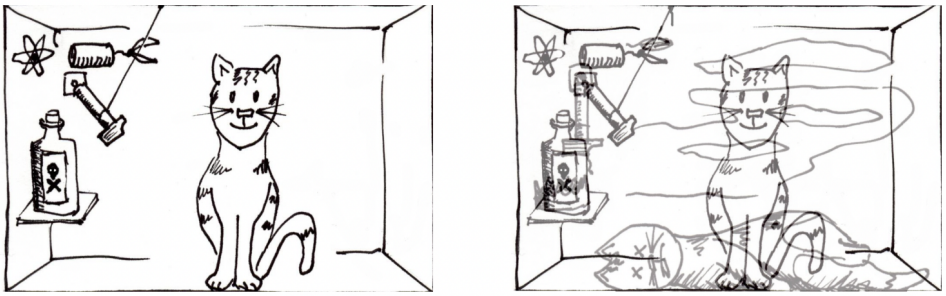


Figure 1: Schrödingers Cat.

However, this cannot be the entire story, because quantum mechanics clearly is not just about our opinions of cats and decaying atoms. In particular, if we prepare an electron in a spin "up" state $|\uparrow\rangle$, then whenever we measure the spin along the z-direction we will find the measurement outcome "up", no matter what we think about electrons and quantum mechanics. So there seems to be some physical property associated with the electron that determines the measurement outcome and is described by the quantum state.

Various interpretations of quantum mechanics attempt to address these (and other) issues. The original interpretation of quantum mechanics was mainly put forward by Niels Bohr, and is called the **Copenhagen interpretation**. Broadly speaking, it says that the quantum state is a convenient fiction, used to calculate the results of measurement outcomes, and that the system cannot be considered separate from the measurement apparatus. Alternatively, there are interpretations of quantum mechanics, such as the **Ghirardi-Rimini-Weber interpretation**, that do ascribe some kind of reality to the state of the system, in which case a physical mechanism for the collapse of the wave function must be given. Many of these interpretations can be classified as hidden variable theories, which postulate that there is a deeper physical reality described by some "hidden variables" that we must average over. This in turn explains the probabilistic nature of quantum mechanics. The problem with such theories is that these hidden variables must be quite weird: they can change instantly depending on events light-years away3 , thus violating Einstein's theory of special relativity. Many physicists do not like this aspect of hidden variable theories.

Alternatively, quantum mechanics can be interpreted in terms of "many worlds": the **Many Worlds interpretation** states that there is one state vector for the entire universe, and that each measurement splits the universe into different branches corresponding to the different measurement outcomes. It is attractive since it seems to be a philosophically consistent interpretation, and while it has been acquiring a growing number of supporters over recent years[4], a lot of physicists have a deep aversion to the idea of parallel universes.

Finally, there is the **epistemic interpretation**, which is very similar to the Copenhagen interpretation in that it treats the quantum state to a large extent as a measure of our knowledge of the quantum system (and the measurement apparatus). At the same time, it denies a deeper underlying reality (i.e., no hidden variables). The attractive feature of this interpretation is that it requires a minimal amount of fuss, and fits naturally with current research in quantum information theory. The downside is that you have to abandon simple scientific realism that allows you to talk about the properties of electrons and photons, and many physicists are not prepared to do that.

As you can see, quantum mechanics forces us to abandon some deeply held (classical) convictions about Nature. Depending on your preference, you may be drawn to one or other interpretation. It is currently not know which interpretation is the correct one.

Exercises

1. Calculate the eigenvalues and the eigenstates of the bit flip operator X, and show that the eigenstates form an orthonormal basis. Calculate the expectation value of X for $|\psi\rangle = 1/\sqrt{3}|0\rangle + i\sqrt{2/3}|1\rangle$.
2. Show that the variance of A vanishes when $|\psi\rangle$ is an eigenstate of A.
3. Prove that an operator is Hermitian if and only if it has real eigenvalues.
4. Show that a qubit in an unknown state $|\psi\rangle$ cannot be copied. This is the no-cloning theorem. Hint: start with a state $|\psi\rangle|i\rangle$ for some initial state $|i\rangle$, and require that for $|\psi\rangle = |0\rangle$ and $|\psi\rangle = |1\rangle$ the cloning procedure is a unitary transformation $|0\rangle|i\rangle \rightarrow |0\rangle|0\rangle$ and $|1\rangle|i\rangle \rightarrow |1\rangle|1\rangle$.
5. The uncertainty principle.
   a. Use the Cauchy-Schwarz inequality to derive the following relation between non-commuting observables A and B:
   $$(\Delta A)^2(\Delta B)^2 \geq \frac{1}{4}|\langle [A,B]\rangle|^2 \tag{2.27}$$
   Hint: define $|f\rangle = (A - \langle A\rangle)|\psi\rangle$ and $|g\rangle = i(B - \langle B\rangle)|\psi\rangle$, and use that $|\langle f \mid g\rangle| \geq \frac{1}{2} \mid \langle f \mid g\rangle + \langle g|f\rangle|$.
   b. Show that this reduces to Heisenberg's uncertainty relation when A and B are canonically conjugate observables, for example position and momentum.
   c. Does this method work for deriving the uncertainty principle between energy and time?
6. Consider the Hamiltonian H and the state $|\psi\rangle$ given by
   $$H = E\begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad |\psi\rangle = \frac{1}{\sqrt{5}}\begin{pmatrix} 1-i \\ 1-i \\ 1 \end{pmatrix} \tag{2.28}$$
   where E is a constant with dimensions of energy. Calculate the energy eigenvalues and the expectation value of the Hamiltonian.
7. Show that the momentum and the total energy can be measured simultaneously only when the potential is constant everywhere. What does a constant potential mean in terms of the dynamics of a particle?

[1]In Hilbert spaces of infinite dimensionality, there are subtle differences between self-adjoint and Hermitian operators. We ignore these subtleties here, because we will be mostly dealing with finite-dimensional spaces.

[2]This is something most people require from a fundamental theory: quantum mechanics should not just break down for macroscopic objects. Indeed, experimental evidence of macroscopic superpositions has been found in the form of "cat states".

[3] . . . even though the averaging over the hidden variables means you can never signal faster than light.

[4]There seems to be some evidence that the Many Worlds interpretation fits well with the latest cosmological models based on string theory

This page titled 2: The Postulates of Quantum Mechanics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

3: Schrödinger and Heisenberg Pictures

So far we have assumed that the quantum states $|\psi(t)\rangle$ describing the system carry the time dependence. However, this is not the only way to keep track of the time evolution. Since all physically observed quantities are expectation values, we can write

$$\begin{aligned}\langle A\rangle &= \mathrm{Tr}[|\psi(t)\rangle\langle\psi(t)|A] = \mathrm{Tr}\big[U(t)|\psi(0)\rangle\langle\psi(0)|U^\dagger(t)A\big] \\ &= \mathrm{Tr}\big[|\psi(0)\rangle\langle\psi(0)|U^\dagger(t)AU(t)\big] \quad \text{(cyclic property)} \\ &\equiv \mathrm{Tr}[|\psi(0)\rangle\langle\psi(0)|A(t)],\end{aligned} \tag{3.1}$$

where we defined the time-varying operator $A(t) = U^\dagger(t)AU(t)$. Clearly, we can keep track of the time evolution in the operators!

- Schrödinger picture: Keep track of the time evolution in the states,
- Heisenberg picture: Keep track of the time evolution in the operators.

We can label the states and operators "S" and "H" depending on the picture. For example,

$$|\psi_H\rangle = |\psi_S(0)\rangle \quad \text{and} \quad A_H(t) = U^\dagger(t)A_SU(t) \tag{3.2}$$

The time evolution for states is given by the Schrödinger equation, so we want a corresponding "Heisenberg equation" for the operators. First, we observe that

$$U(t) = \exp\left(-\frac{i}{\hbar}Ht\right), \tag{3.3}$$

such that

$$\frac{d}{dt}U(t) = -\frac{i}{\hbar}HU(t) \tag{3.4}$$

Next, we calculate the time derivative of $\langle A\rangle$:

$$\frac{d}{dt}\mathrm{Tr}[|\psi_S(t)\rangle\langle\psi_S(t)|A_S] = \frac{d}{dt}\langle\psi_S(t)|A_S|\psi_S(t)\rangle = \frac{d}{dt}\langle\psi_H|A_H(t)|\psi_H\rangle \tag{3.5}$$

The last equation follows from Eq. (3.1). We can now calculate

$$\begin{aligned}\frac{d}{dt}\langle\psi_S(t)|A_S|\psi_S(t)\rangle &= \frac{d}{dt}\langle\psi_S(0)\big|U^\dagger(t)A_SU(t)\big|\psi_S(0)\rangle \\ &= \Big\langle\psi_S(0)\Big|\Big[\dot{U}^\dagger(t)A_SU(t) + U^\dagger(t)\dot{A}_SU(t) + U^\dagger(t)A_S\dot{U}(t)\Big]\Big|\psi_S(0)\Big\rangle \\ &= \left\langle\psi_H\left|\left[\frac{i}{\hbar}HA_H(t) - \frac{i}{\hbar}A_H(t)H + \frac{\partial A_H(t)}{\partial t}\right]\right|\psi_H\right\rangle \\ &= -\frac{i}{\hbar}\langle\psi_H|[A_H(t),H]|\psi_H\rangle + \left\langle\psi_H\left|\frac{\partial A_H(t)}{\partial t}\right|\psi_H\right\rangle \\ &= \left\langle\psi_H\left|\frac{dA_H(t)}{dt}\right|\psi_H\right\rangle\end{aligned} \tag{3.6}$$

Since this must be true for all $|\psi_H\rangle$, this is an operator identity:

$$\frac{dA_H(t)}{dt} = -\frac{i}{\hbar}[A_H(t),H] + \frac{\partial A_H(t)}{\partial t} \tag{3.7}$$

This is the Heisenberg equation. Note the difference between the "straight d" and the "curly ∂" in the time derivative and the partial time derivative, respectively. The partial derivative deals only with the explicit time dependence of the operator. In many cases (such as position and momentum) this is zero.

We have seen that both the Schrödinger and the Heisenberg equation follows from Von Neumann's Hilbert space formalism of quantum mechanics. Consequently, we have proved that this formalism properly unifies both Schrödingers wave mechanics, and Heisenberg, Born, and Jordans matrix mechanics.

As an example, consider a qubit with time evolution determined by the Hamiltonian $H = \frac{1}{2}\hbar\omega Z$, with $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. This may be a spin in a magnetic field, for example, such that $\omega = -eB/mc$. We want to calculate the time evolution of the operator $X_H(t)$. Since we work in the Heisenberg picture alone, we will omit the subscript H. First, we evaluate the commutator in the Heisenberg equation

$$i\hbar\frac{1}{2}\frac{dX}{dt} = \frac{1}{2}[X, H] = -i\hbar\frac{\omega}{2}Y, \tag{3.8}$$

where we defined $Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$. So now we must know the time evolution of Y as well:

$$i\hbar\frac{1}{2}\frac{dY}{dt} = \frac{1}{2}[Y, H] = i\hbar\frac{\omega}{2}X \tag{3.9}$$

These are two coupled linear equations, which are relatively easy to solve:

$$\dot{X} = -\omega Y \quad \text{and} \quad \dot{Y} = \omega X \quad \text{and} \quad \dot{Z} = 0 \tag{3.10}$$

We can define two new operators $S_\pm = X \pm iY$, and obtain

$$\dot{S}_\pm = -\omega Y \pm i\omega X = \pm i\omega S_\pm. \tag{3.11}$$

Solving these two equations yields $S_\pm(t) = S_\pm(0)e^{\pm i\omega t}$, and this leads to

$$\begin{aligned} X(t) &= \frac{S_+(t) + S_-(t)}{2} = \frac{S_+(0)e^{i\omega t} + S_-(0)e^{-i\omega t}}{2} \\ &= \frac{1}{2}\left[X(0)e^{i\omega t} + iY(0)e^{i\omega t} + X(0)e^{-i\omega t} - iY(0)e^{-i\omega t}\right] \\ &= X(0)\cos(\omega t) - Y(0)\sin(\omega t). \end{aligned} \tag{3.12}$$

You are asked to show that $Y(t) = Y(0)\cos(\omega t) + X(0)\sin(\omega t)$ in exercise 3.

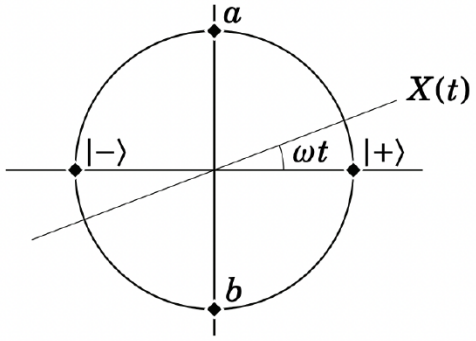
We now take $|\psi_H\rangle = |0\rangle$ and $X(0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $Y(0) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$. The expectation value of $X(t)$ is then readily calculated to be

$$\langle 0|X(t)|0\rangle = \cos(\omega t)\langle 0|X(0)|0\rangle - \sin(\omega t)\langle 0|Y(0)|0\rangle = 0 \tag{3.13}$$

Alternatively, when $|\psi_H\rangle = |\pm\rangle$, we find

$$\langle +|X(t)|+\rangle = \cos(\omega t) \quad \text{and} \quad \langle +|Y(t)|+\rangle = \sin(\omega t) \tag{3.14}$$

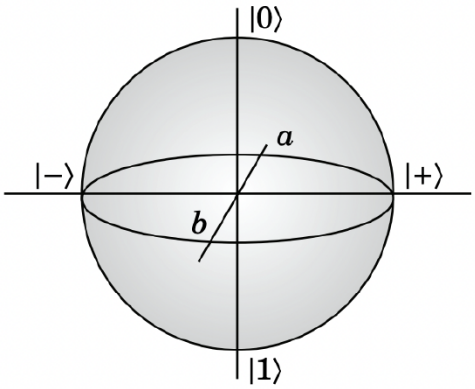
This is a circular motion in time:



The eigenstate of $X(\pi/2)$ is point a, and the eigenstate of $X(-\pi/2)$ is point b. Furthermore, $X(\pm\pi/2) = \mp Y(0)$, and the states at point a and b are therefore the eigenstates of Y:

$$|\psi_a\rangle = \frac{|0\rangle - i|1\rangle}{\sqrt{2}} \quad \text{and} \quad |\psi_b\rangle = \frac{|0\rangle + i|1\rangle}{\sqrt{2}} \tag{3.15}$$

A natural question to ask is where the states $|0\rangle$ and $|1\rangle$ fit in this picture. These are the eigenstates of the operator Z, which we used to generate the unitary time evolution. Clearly the states on the circle never become either $|0\rangle$ or $|1\rangle$, so we need to add another dimension:



This is called the Bloch sphere, and operators are represented by straight lines through the origin. The axis of rotation for the straight lines that rotate with time is determined by the Hamiltonian. In the above case the Hamiltonian was proportional to Z, which means that the straight lines rotate around the axis through the eigensates of Z, which are $|0\rangle$ and $|1\rangle$.

? Exercises

1. Show that for the Hamiltonian $H_S = H_H$.
2. The harmonic oscillator has the energy eigenvalue equation $H|n\rangle = \hbar\omega\left(n+\frac{1}{2}\right)|n\rangle$.
   a. The classical solution of the harmonic oscillator is given by

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle, \tag{3.16}$$

   in the limit of $|\alpha| \gg 1$. Show that $|\alpha\rangle$ is a properly normalized state for any $\alpha \in \mathbb{C}$.
   b. Calculate the time-evolved state $|\alpha(t)\rangle$.
   c. We introduce the ladder operators $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$ and $\hat{a}^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle$. Show that the number operator defined by $\hat{n}|n\rangle = n|n\rangle$ can be written as $\hat{n} = \hat{a}^\dagger\hat{a}$.
   d. Write the coherent state $|\alpha\rangle$ as a superposition of ladder operators acting on the ground state $|0\rangle$.
   e. Note that the ground state is time-independent $(U(t)|0\rangle = |0\rangle)$. Calculate the time evolution of the ladder operators.
   f. Calculate the position $\hat{q} = \left(\hat{a}+\hat{a}^\dagger\right)/2$ and momentum $\hat{p} = -i\left(\hat{a}-\hat{a}^\dagger\right)/2$ of the harmonic oscillator in the Heisenberg picture. Can you identify the classical harmonic motion?
3. Let A be an operator given by $A = a_0\mathbb{I} + a_xX + a_yY + a_zZ$. Calculate the matrix $A(t)$ given the Hamiltonian $H = \frac{1}{2}\hbar\omega Z$, and show that A is Hermitian when the a_μ are real.
4. The interaction picture.
   a. Let the Hamiltonian of a system be given by $H = H_0 + V$, with $H_0 = p^2/2m$. Using $|\psi(t)\rangle_I = U_0^\dagger(t)|\psi(t)\rangle_S$ with $U_0(t) = \exp(-iH_0t/\hbar)$, calculate the time dependence of an operator in the interaction picture $A_I(t)$.
   b. Defining $H_I(t) = U_0^\dagger(t)VU_0(t)$, show that

$$i\hbar\frac{d}{dt}|\psi(t)\rangle_I = H_I(t)|\psi(t)\rangle_I \tag{3.17}$$

   Is H_I identical to H_H and H_S?
5. The time operator in quantum mechanics.
   a. Let $H|\psi\rangle = E|\psi\rangle$, and assume the existence of a time operator conjugate to H, i.e., $[H,T] = i\hbar$. Show that

$$He^{i\omega T}|\psi\rangle = (E-\hbar\omega)e^{i\omega T}|\psi\rangle \tag{3.18}$$

   b. Given that $\omega \in \mathbb{R}$, calculate the spectrum of H.

c. The energy of a system must be bounded from below in order to avoid infinite decay to ever lower energy states. What does this mean for T?

6. Consider a three-level atom with two (degenerate) low-lying states $|0\rangle$ and $|1\rangle$ with zero energy, and a high level $|e\rangle$ (the "excited" state) with energy $\hbar\omega$. The low levels are coupled to the excited level by optical fields $\Omega_0 \cos\omega_0 t$ and $\Omega_1 \cos\omega_1 t$, respectively.
   a. Give the (time-dependent) Hamiltonian H for the system.
   b. The time dependence in H is difficult to deal with, so we must transform to the rotating frame via some unitary transformation $U(t)$. Show that
   $$H' = U(t)HU^\dagger(t) - i\hbar U\frac{dU^\dagger}{dt}$$
   You can use the Schrödinger equation with $|\psi\rangle = U^\dagger|\psi'\rangle$.
   c. Calculate H' if $U(t)$ is given by
   $$U(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-i(\omega_0-\omega_1)t} & 0 \\ 0 & 0 & e^{-i\omega_0 t} \end{pmatrix}$$
   Why can we ignore the remaining time dependence in H'? This is called the Rotating Wave Approximation.
   d. Calculate the $\lambda = 0$ eigenstate of H' in the case where $\omega_0 = \omega_1$.
   e. Design a way to bring the atom from the state $|0\rangle$ to $|1\rangle$ without ever populating the state $|e\rangle$.

This page titled 3: Schrödinger and Heisenberg Pictures is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

4: Mixed States and the Density Operator

So far we have considered states as vectors in Hilbert space that, according to the first postulate, contain all the information about the system. In reality, however, we very rarely have complete information about a system. For example, the system may have interacted with its environment, which introduces some uncertainty in our knowledge of the state of the system. The question is therefore how we describe systems given incomplete information. Much of contemporary research in quantum mechanics is about gaining full control over the quantum system (meaning to minimize the interaction with the environment). This includes the field of quantum information and computation. The concept of incomplete information is therefore central to modern quantum mechanics.

4.1: Mixed States

4.2: Decoherence

4.3: Imperfect Measurements

This page titled 4: Mixed States and the Density Operator is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

4.1: Mixed States

First, we recall some properties of the trace:

- $\mathrm{Tr}(aA) = a\,\mathrm{Tr}(A)$,
- $\mathrm{Tr}(A+B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$.

Also remember that we can write the expectation value of A as

$$\langle A \rangle = \mathrm{Tr}(|\psi\rangle\langle\psi|A), \tag{4.1}$$

where $|\psi\rangle$ is the state of the system. It tells us everything there is to know about the system. But what if we don't know everything?

As an example, consider that Alice prepares a qubit in the state $|0\rangle$ or in the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ depending on the outcome of a balanced (50:50) coin toss. How does Bob describe the state before any measurement? First, we cannot say that the state is $\frac{1}{2}|0\rangle + \frac{1}{2}|+\rangle$, because this is not normalized!

The key to the solution is to observe that the expectation values must behave correctly. The expectation value $\langle A \rangle$ is the average of the eigenvalues of A for a given state. If the state is itself a statistical mixture (as in the example above), then the expectation values must also be averaged. So for the example, we require that for any A

$$\begin{aligned} \langle A \rangle &= \frac{1}{2}\langle A \rangle_0 + \frac{1}{2}\langle A \rangle_+ = \frac{1}{2}\mathrm{Tr}(|0\rangle\langle 0|A) + \frac{1}{2}\mathrm{Tr}(|+\rangle\langle +|A) \\ &= \mathrm{Tr}\left[\left(\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +|\right) A\right] \\ &\equiv \mathrm{Tr}(\rho A), \end{aligned} \tag{4.2}$$

where we defined

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +|. \tag{4.3}$$

The statistical mixture is therefore properly described by an operator, rather than a simple vector. We can generalize this as

$$\rho = \sum_k p_k \left|\psi_k\right\rangle \left\langle\psi_k\right|, \tag{4.4}$$

where the p_k are probabilities that sum up to one ($\sum_k p_k = 1$) and the $|\psi_k\rangle$ are normalized states (not necessarily complete or orthogonal). Since ρ acts as a weight, or a density, in the expectation value, we call it the density operator. We can diagonalize ρ to find the spectral decomposition

$$\rho = \sum_j \lambda_j \left|\lambda_j\right\rangle \left\langle\lambda_j\right|, \tag{4.5}$$

where $\{|\lambda_j\rangle\}_j$ forms a complete orthonormal basis, $0 \le \lambda_j \le 1$, and $\sum_j \lambda_j = 1$. We can also show that ρ is a positive operator:

$$\begin{aligned} \langle\psi|\rho|\psi\rangle &= \sum_{jk} c_j^* c_k \left\langle\lambda_j|\rho|\lambda_k\right\rangle = \sum_{jk} c_j^* c_k \left\langle \lambda_j \middle| \sum_l \lambda_l \middle| \lambda_l \right\rangle \left\langle \lambda_l \mid \lambda_k \right\rangle \\ &= \sum_{jkl} c_j^* c_k \lambda_l \left\langle \lambda_j \mid \lambda_l \right\rangle \left\langle \lambda_l \mid \lambda_k \right\rangle = \sum_{jkl} c_j^* c_k \lambda_l \delta_{jl} \delta_{lk} \\ &= \sum_l \lambda_l |c_l|^2 \\ &\ge 0 \end{aligned} \tag{4.6}$$

In general, an operator ρ is a valid density operator if and only if it has the following three properties:

1. $\rho^\dagger = \rho$,
2. $\mathrm{Tr}(\rho) = 1$,
3. $\rho \ge 0$.

The density operator is a generalization of the state of a quantum system when we have incomplete information. In the special case where one of the $p_j = 1$ and the others are zero, the density operator becomes the projector $|\psi_j\rangle\langle\psi_j|$. In other words, it is

completely determined by the state vector $|\psi_j\rangle$. We call these pure states. The statistical mixture of pure states giving rise to the density operator is called a **mixed state**.

The unitary evolution of the density operator can be derived directly from the Schrödinger equation $i\hbar\partial_t|\psi\rangle = H|\psi\rangle$:

$$\begin{aligned} i\hbar\frac{d\rho}{dt} &= i\hbar\frac{d}{dt}\sum_j p_j\,|\psi_j\rangle\,\langle\psi_j| \\ &= i\hbar\sum_j\left\{\frac{dp_j}{dt}|\psi_j\rangle\,\langle\psi_j| + p_j\left[\left(\frac{d}{dt}|\psi_j\rangle\right)\langle\psi_j| + |\psi_j\rangle\left(\frac{d}{dt}\langle\psi_j|\right)\right]\right\} \\ &= i\hbar\frac{\partial\rho}{\partial t} + H\rho - \rho H \\ &= [H,\rho] + i\hbar\frac{\partial\rho}{\partial t} \end{aligned} \tag{4.7}$$

This agrees with the Heisenberg equation for operators, and it is sometimes known as the Von Neumann equation. In most problems the probabilities p_j have no explicit time-dependence, and $\partial_t\rho = 0$.

This page titled 4.1: Mixed States is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

4.2: Decoherence

The density operator allows us to consider the phenomenon of **decoherence**. Consider the pure state $|+\rangle$. In matrix notation with respect to the basis $\{|+\rangle, |-\rangle\}$, this can be written as

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{4.8}$$

The trace is 1, and one of the eigenvalues is 1, as required for a pure state. We can also write the density operator in the basis $\{|0\rangle, |1\rangle\}$:

$$\rho = |+\rangle\langle+| = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \times \frac{1}{\sqrt{2}}(1,1) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{4.9}$$

Notice how the outer product (as opposed to the inner product) of two vectors creates a matrix representation of the corresponding projection operator.

Let the time evolution of $|+\rangle$ be given by

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \rightarrow \frac{|0\rangle + e^{i\omega t}|1\rangle}{\sqrt{2}}. \tag{4.10}$$

The corresponding density operator becomes

$$\rho(t) = \frac{1}{2}\begin{pmatrix} 1 & e^{i\omega t} \\ e^{-i\omega t} & 1 \end{pmatrix}. \tag{4.11}$$

The "population" in the state $|+\rangle$ is given by the expectation value

$$\langle+|\rho(t)|+\rangle = \frac{1}{2} + \frac{1}{2}\cos(\omega t). \tag{4.12}$$

This oscillation is due to the off-diagonal elements of $\rho(t)$, and it is called the coherence of the system (see Fig. 2).

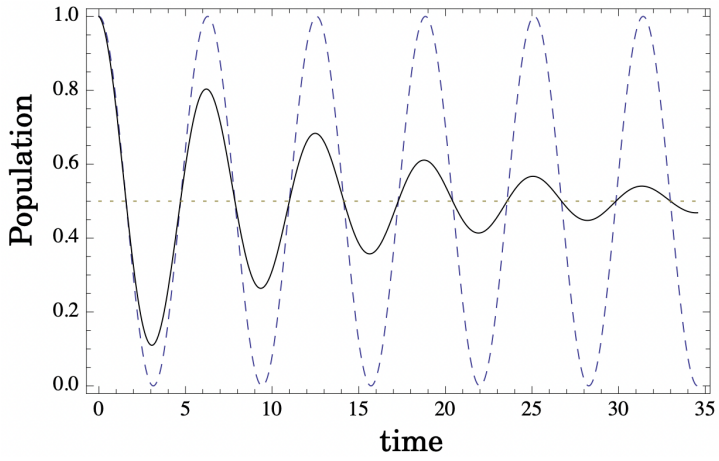


Figure 2: Population in the state $|+\rangle$ with decoherence (solid curve) and without (dashed curve).

The state is pure at any time t. In real physical systems the coherence often decays exponentially at a rate γ, and the density matrix can be written as

$$\rho(t) = \frac{1}{2}\begin{pmatrix} 1 & e^{i\omega t - \gamma t} \\ e^{-i\omega t - \gamma t} & 1 \end{pmatrix}. \tag{4.13}$$

The population in the state $|+\rangle$ decays accordingly as

$$\langle+|\rho(t)|+\rangle = \frac{1}{2} + \frac{e^{-\gamma t}\cos(\omega t)}{2}. \tag{4.14}$$

This is called decoherence of the system, and the value of γ depends on the physical mechanism that leads to the decoherence.

The decoherence described above is just one particular type, and is called **dephasing**. Another important decoherence mechanism is relaxation to the ground state. If the state $|1\rangle$ has a larger energy than $|0\rangle$ there may be processes such as spontaneous emission that drive the system to the ground state. Combining these two decay processes, we can write the density operator as

$$\rho(t)=\frac{1}{2}\begin{pmatrix}2-e^{-\gamma_1 t} & e^{i\omega t-\gamma_2 t}\\ e^{-i\omega t-\gamma_2 t} & e^{-\gamma_1 t}\end{pmatrix}. \tag{4.15}$$

The study of decoherence is currently one of the most important research areas in quantum physics.

This page titled 4.2: Decoherence is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

4.3: Imperfect Measurements

Postulates 3 and 5 determine what happens when a quantum system is subjected to a measurement. In particular, these postulates concern ideal measurements. In practice, however, we often have to deal with imperfect measurements that include noise. As a simple example, consider a photodetector: not every photon that hits the detector will result in a detector "click", which tells us that there indeed was a photon. How can we describe situations like these?

First, let's recap ideal measurements. We can ask the question what will be the outcome of a single measurement of the observable A. We know from Postulate 3 that the outcome m must be an eigenvalue a_m of A. If the spectral decomposition of A is given by

$$A = \sum_m a_m |m\rangle\langle m|, \tag{4.16}$$

then the probability of finding measurement outcome m is given by the Born rule

$$p(m) = |\langle m \mid \psi\rangle|^2 = \mathrm{Tr}(P_m|\psi\rangle\langle\psi|) \equiv \langle P_m\rangle\,, \tag{4.17}$$

where we introduced the operator $P_m = |m\rangle\langle m|$. One key interpretation of $p(m)$ is as the expectation value of the operator P_m associated with measurement outcome m.

When a measurement does not destroy the system, the state of the system must be updated after the measurement to reflect the fact that another measurement of A immediately following the first will yield outcome m with certainty. The update rule is given by Eq. (2.24)

$$|\psi\rangle \to \frac{P_m|\psi\rangle}{\sqrt{\left\langle\psi\left|P_m^\dagger P_m\right|\psi\right\rangle}} \tag{4.18}$$

where the square root in the denominator is included to ensure proper normalization of the state after the measurement. However, this form is not so easily generalized to measurements yielding incomplete information, so instead we will write (using $P_m^2 = P_m = P_m^\dagger$)

$$|\psi\rangle\langle\psi| \to \frac{P_m|\psi\rangle\langle\psi|P_m^\dagger}{\mathrm{Tr}\left(P_m|\psi\rangle\langle\psi|P_m^\dagger\right)} \tag{4.19}$$

From a physical perspective this notation is preferable, since the unobservable global phase of $\langle m \mid \psi\rangle$ no longer enters the description. Up to this point, both state preparation and measurement were assumed to be ideal. How must this formalism of ideal measurements be modified in order to take into account measurements that yield only partial information about the system? First, we must find the probability for measurement outcome m, and second, we have to formulate the update rule for the state after the measurement.

When we talk about imperfect measurements, we must have some way of knowing (or suspecting) how the measurement apparatus fails. For example, we may suspect that a photon hitting a photodetector has only a finite probability of triggering a detector "click". We therefore describe our measurement device with a general probability distribution $\{q_j(m)\}$, where $q_j(m)$ is the probability that the measurement outcome m in the detector is triggered by a system in the state $|\psi_j\rangle$. The probabilities $q_j(m)$ can be found by modelling the physical aspects of the measurement apparatus. The accuracy of this model can then be determined by experiment. The probability of measurement outcome m for the ideal case is given in Eq. (4.17) by $\langle P_m\rangle$. When the detector is not ideal, the probability of finding outcome m is given by the weighted average over all possible expectation values $\langle P_j\rangle$ that can lead to m:

$$\begin{aligned} p(m) &= \sum_j q_j(m)\,\langle P_j\rangle = \sum_j q_j(m)\,\mathrm{Tr}(P_j|\psi\rangle\langle\psi|) \\ &= \mathrm{Tr}\left[\sum_j q_j(m) P_j|\psi\rangle\langle\psi|\right] \equiv \mathrm{Tr}(E_m|\psi\rangle\langle\psi|)\,, \\ &= \langle E_m\rangle \end{aligned} \tag{4.20}$$

where we defined the operator $\boldsymbol{E}_m$ associated with outcome m as

$$E_m = \sum_j q_j(m) P_j \tag{4.21}$$

Each possible measurement outcome m has an associated operator E_m, the expectation value of which is the probability of obtaining m in the measurement. The set of $\boldsymbol{E}_m$ is called a Positive Operator-Valued Measure (POVM). While the total number of ideal measurement outcomes, modelled by P_m, must be identical to the dimension of the Hilbert space (ignoring the technical complication of degeneracy), this is no longer the case for the POVM described by E_m; there can be more or fewer measurement outcomes, depending on the physical details of the measurement apparatus. For example, the measurement of an electron spin in a Stern-Gerlach apparatus can have outcomes "up", "down", or "failed measurement". The first two are determined by the position of the fluorescence spot on the screen, and the last may be the situation where the electron fails to produce a spot on the screen. Here, the number of possible measurement outcomes is greater than the number of eigenstates of the spin operator. Similarly, photodetectors in Geiger mode have only two possible outcomes, namely a "click" or "no click" depending on whether the detector registered photons or not. However, the number of eigenstates of the intensity operator (the photon number states) is infinite.

Similar to the density operator, the POVM elements E_m have three key properties. First, the $p(m)$ are probabilities and therefore real, so for all states $|\psi\rangle$ we have

$$\langle E_m \rangle^* = \langle E_m \rangle \quad \Longleftrightarrow \langle \psi | E_m | \psi \rangle = \left\langle \psi \middle| E_m^\dagger \middle| \psi \right\rangle, \tag{4.22}$$

and E_m is therefore Hermitian. Second, since $\sum_m p(m) = 1$ we have

$$\sum_m p(m) = \sum_m \langle \psi | E_m | \psi \rangle = \left\langle \psi \middle| \sum_m E_m \middle| \psi \right\rangle = 1, \tag{4.23}$$

for all $|\psi\rangle$, and therefore

$$\sum_m E_m = \mathbb{I}, \tag{4.24}$$

where $\mathbb{I}$ is the identity operator. Finally, since $p(m) = \langle \psi | E_m | \psi \rangle \geq 0$ for all $|\psi\rangle$, the POVM element $\boldsymbol{E}_m$ is a positive operator. Note the close analogy of the properties of the POVM and the density operator. In particular, just as in the case of density operators, POVMs are defined by these three properties.

The second question about generalized measurements is how the measurement outcomes should be used to update the quantum state of the system. The rule for ideal von Neumann measurements is given in Eq. (4.19), which can be generalized immediately to density operators using the techniques presented above. This yields

$$\rho \to \frac{P_m \rho P_m^\dagger}{\mathrm{Tr}\left(P_m \rho P_m^\dagger\right)} \tag{4.25}$$

What if we have instead a measurement apparatus described by a POVM? Consistency with the Born rule in Eq. (4.20) requires that we again replace P_m, associated with measurement outcome m, with a probability distribution over all P_j:

$$\rho \to \sum_j q_j(m) \frac{P_j \rho P_j^\dagger}{\mathrm{Tr}\left[\sum_k q_k(m) P_k \rho P_k^\dagger\right]} = \sum_j q_j(m) \frac{P_j \rho P_j^\dagger}{\mathrm{Tr}(E_m \rho)}, \tag{4.26}$$

where we adjusted the normalization factor to maintain $\mathrm{Tr}(\rho) = 1$. We also used that the POVM element $\boldsymbol{E}_m$ in Eq. (4.21) can be written as

$$E_m = \sum_j q_j(m) P_j^\dagger P_j \tag{4.27}$$

If we rescale the P_j by a factor $\sqrt{q_j(m)}$, we obtain the standard form of the POVM

$$E_m = \sum_j A_{jm}^\dagger A_{jm} \tag{4.28}$$

where $A_{jm} = \sqrt{q_j(m)} P_j$ are the so-called the Kraus operators. Consequently, the update rule can be written as

$$\rho \to \frac{\sum_j A_{jm}\rho A_{jm}^\dagger}{\mathrm{Tr}(E_m\rho)} \tag{4.29}$$

which generally yields a mixed state (described by a density operator) after an incomplete measurement.

Finally, we can consider the case of some nonunitary evolution without a measurement (e.g., when the system interacts with its environment). We can model this purely formally by removing the index of the measurement outcome m from the description (since there are no measurement outcomes). The most general evolution then takes the form

$$\rho \to \rho' = \sum_k A_k \rho A_k^\dagger, \tag{4.30}$$

and the question is what form the Kraus operators A_k take. We will return to this in section 6.

? Exercises

1. The density matrix.
   a. Show that $\frac{1}{2}|0\rangle + \frac{1}{2}|+\rangle$ is not a properly normalized state.
   b. Show that $\mathrm{Tr}(\rho) = 1$ with ρ given by Eq. (4.3), and then prove that any density operator has unit trace and is Hermitian.
   c. Show that density operators are convex, i.e., that $\rho = w_1\rho_1 + w_2\rho_2$ with $w_1 + w_2 = 1 (w_1, w_2 \geq 0)$, and ρ_1, ρ_2 again density operators.
   d. Calculate the expectation value of A using the two representations of ρ in terms of p_i and the spectral decomposition. What is the difference in the physical interpretation of p_j and λ_j?
2. Using the identity $\langle x|A|\psi\rangle = A\psi(x)$, and the resolution of the identity $\int dx|x\rangle\langle x| = \mathbb{I}$, calculate the expectation value for an operator A, given a mixed state of wave functions.
3. Calculate P^2 with P given by

$$P = \begin{pmatrix} a \\ b \end{pmatrix} (a^*, b^*) = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix} \quad \text{and} \quad |a|^2 + |b|^2 = 1 \tag{4.31}$$

4. Calculate the eigenvalues of the density matrix in Eq. (4.15), and show that $\gamma_1 \leq 2\gamma_2$. Hint: it is difficult to derive the inequality directly, so you should try to demonstrate that $\gamma_1 \geq 2\gamma_2$ leads to a contradiction.
5. A system with energy eigenstates $|E_n\rangle$ is in thermal equilibrium with a heat bath at temperature T. The probability for the system to be in state E_n is proportional to $e^{-E_n/kT}$.
   a. Write the Hamiltonian of the system in terms of $\boldsymbol{E}_n$ and $|E_n\rangle$.
   b. Give the normalized density operator ρ for the system as a function of the Hamiltonian.
   c. We identify the normalization of ρ with the partition function $\mathscr{Z}$. Calculate the average energy directly and via

$$\langle E\rangle = -\frac{\partial \ln \mathscr{Z}}{\partial \beta} \tag{4.32}$$

   of the system, where $\beta = 1/(kT)$.
   d. What is the entropy $S = k\ln\mathscr{Z}$ if the system is a harmonic oscillator? Comment on the limit $T \to 0$.
6. Lossy photodetectors.
   a. The state of a beam of light can be written in the photon number basis $|n\rangle$ as $|\psi\rangle = \sum_n c_n|n\rangle$. What are the possible measurement outcomes for a perfect photon detector? Calculate the probabilities of the measurement outcome using projection operators.
   b. Suppose that the detector can only tell the difference between the presence and absence of photons (a so-called "bucket detector"). How do we calculate the probability of finding the measurement outcomes?
   c. Real bucket detectors have a finite efficiency η, which means that each photon has a probability η of triggering the detector. Calculate the probabilities of the measurement outcomes.
   d. What other possible imperfections do realistic bucket detectors have?
7. An electron with spin state $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$ and $|\alpha|^2 + |\beta|^2 = 1$ is sent through a SternGerlach apparatus to measure the spin in the z direction (i.e., the $|\uparrow\rangle, |\downarrow\rangle$ basis).

a. What are the possible measurement outcomes? If the position of the electron is measured by an induction loop rather than a screen, what is the state of the electron immediately after the measurement?
b. Suppose that with probability p the induction loops fail to give a current signifying the presence of an electron. What are the three possible measurement outcomes? Give the POVM.
c. Calculate the probabilities of the measurement outcomes, and the state of the electron imediately after the measurement (for all possible outcomes).

This page titled 4.3: Imperfect Measurements is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

5: Composite Systems and Entanglement

5.1: Composite Systems

5.2: Entanglement

5.3: Quantum teleportation

This page titled 5: Composite Systems and Entanglement is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

5.1: Composite Systems

Suppose we have two systems, described by Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$, respectively. We can choose orthonormal bases for each system:

$$\mathscr{H}_1 : \{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_N\rangle\} \quad \text{and } \mathscr{H}_2 : \{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_M\rangle\} \tag{5.1}$$

The respective dimensions of $\mathscr{H}_1$ and $\mathscr{H}_2$ are N and M. We can construct $N \times M$ basis states for the composite system via $|\phi_j\rangle$ and $|\psi_k\rangle$. This implies that the total Hilbert space of the composite system can be spanned by the tensor product

$$\{|\phi_j\rangle \otimes |\psi_k\rangle\}_{jk} \quad \text{on} \quad \mathscr{H}_{1+2} = \mathscr{H}_1 \otimes \mathscr{H}_2 \tag{5.2}$$

An arbitrary pure state on $\mathscr{H}_{1+2}$ can be written as

$$|\Psi\rangle = \sum_{jk} c_{jk} |\phi_j\rangle \otimes |\psi_k\rangle \equiv \sum_{jk} c_{jk} |\phi_j, \psi_k\rangle \tag{5.3}$$

For example, the system of two qubits can be written on the basis $\{|0,0\rangle, |0,1\rangle, |1,0\rangle, |1,1\rangle\}$If system 1 is in state $|\phi\rangle$ and system 2 is in state $|\psi\rangle$, the partial trace over system 2 yields

$$\mathrm{Tr}_2(|\phi,\psi\rangle\langle\phi,\psi|) = \mathrm{Tr}_2(|\phi\rangle\langle\phi| \otimes |\psi\rangle\langle\psi|) = |\phi\rangle\langle\phi| \,\mathrm{Tr}(|\psi\rangle\langle\psi|) = |\phi\rangle\langle\phi|, \tag{5.4}$$

since the trace over any density operator is 1. We have now lost system 2 from our description! Therefore, taking the partial trace without inserting any other operators is the mathematical version of forgetting about it. This is a very useful feature: you often do not want to deal with every possible system you are interested in. For example, if system 1 is a qubit, and system two is a very large environment the partial trace allows you to "trace out the environment".

However, tracing out the environment will not always leave you with a pure state as in Eq. (5.4). If the system has interacted with the environment, taking the partial trace generally leaves you with a mixed state. This is due to entanglement between the system and its environment.

This page titled 5.1: Composite Systems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

5.2: Entanglement

Consider the following experiment: Alice and Bob each blindly draw a marble from a vase that contains one black and one white marble. Let's call the state of the write marble $|0\rangle$ and the state of the black marble $|1\rangle$. If we describe this classical experiment quantum mechanically (we can always do this, because classical physics is contained in quantum physics), then there are two possible states, $|1,0\rangle$ and $|0,1\rangle$. Since blind drawing is a statistical procedure, the state of the marbles held by Alice and Bob is the mixed state

$$\rho = \frac{1}{2}|0,1\rangle\left\langle 0,1\left|+\frac{1}{2}\right|1,0\right\rangle\langle 1,0| \tag{5.5}$$

From Alice's perspective, the state of her marble is obtained by tracing over Bob's marble:

$$\rho_A = \mathrm{Tr}_B(\rho) = \frac{1}{2}|0\rangle\left\langle 0\left|+\frac{1}{2}\right|1\right\rangle\langle 1| \tag{5.6}$$

This is what we expect: Alice has a 50:50 probability of finding "white" or "black" when she looks at her marble (i.e., when she measures the colour of the marble).

Next, consider what the state of Bob's marble is when Alice finds a white marble. Just from the setup we know that Bob's marble must be black, because there was only one white and one black marble in the vase. Let's see if we can reproduce this in our quantum mechanical description. Finding a white marble can be described mathematically by a projection operator $|0\rangle\langle 0|$ (see Eq. (2.24)). We need to include this operator in the trace over Alice's marble's Hilbert space:

$$\rho_B = \frac{\mathrm{Tr}_A(|0\rangle_A\langle 0|\rho)}{\mathrm{Tr}(|0\rangle_A\langle 0|\rho)} = |1\rangle\langle 1|, \tag{5.7}$$

which we set out to prove: if Alice finds that when she sees that her marble is white, she describes the state of Bob's marble as black. Based on the setup of this experiment, Alice knows instantaneously what the state of Bob's marble is as soon as she looks at her own marble. There is nothing spooky about this; it just shows that the marbles held by Alice and Bob are correlated.

Next, consider a second experiment: By some procedure, the details of which are not important right now, Alice and Bob each hold a two-level system (a qubit) in the pure state

$$|\Psi\rangle_{AB} = \frac{|0,1\rangle + |1,0\rangle}{\sqrt{2}} \tag{5.8}$$

Since $|1,0\rangle$ and $|0,1\rangle$ are valid quantum states, by virtue of the first postulate of quantum mechanics $|\Psi\rangle_{AB}$ is also a valid quantum mechanical state. It is not difficult to see that these systems are also correlated in the states $|0\rangle$ and $|1\rangle$: When Alice finds the value "0", Bob must find the value "1", and vice versa. We can write this state as a density operator

$$\begin{aligned}\rho &= \frac{1}{2}(|0,1\rangle + |1,0\rangle)(\langle 0,1| + \langle 1,0|) \\ &= \frac{1}{2}(|0,1\rangle\langle 0,1| + |0,1\rangle\langle 1,0| + |1,0\rangle\langle 0,1| + |1,0\rangle\langle 1,0|).\end{aligned} \tag{5.9}$$

Notice the two extra terms with respect to Eq. (5.5). If Alice now traces out Bob's system, she finds that the state of her marble is

$$\rho_A = \mathrm{Tr}_B(\rho) = \frac{1}{2}|0\rangle\left\langle 0\left|+\frac{1}{2}\right|1\right\rangle\langle 1|. \tag{5.10}$$

In other words, even though the total system was in a pure state, the subsystem held by Alice (and Bob, check this) is mixed! We can try to put the two states back together:

$$\begin{aligned}\rho_A \otimes \rho_B &= \left(\frac{1}{2}|0\rangle\left\langle 0\left|+\frac{1}{2}\right|1\right\rangle\langle 1|\right) \otimes \left(\frac{1}{2}|0\rangle\left\langle 0\left|+\frac{1}{2}\right|1\right\rangle\langle 1|\right) \\ &= \frac{1}{4}(|0,0\rangle\langle 0,0| + |0,1\rangle\langle 0,1| + |1,0\rangle\langle 1,0| + |1,1\rangle\langle 1,1|)\end{aligned}, \tag{5.11}$$

but this is not the state we started out with! It is also a mixed state, instead of the pure state we started with. Since mixed states mean incomplete knowledge, there must be some information in the combined system that does not reside in the subsystems alone!

This is called entanglement.

Entanglement arises because states like $(|0,1\rangle+|1,0\rangle)/\sqrt{2}$ cannot be written as a tensor product of two pure states $|\psi\rangle\otimes|\phi\rangle$. These latter states are called separable. In general a state is separable if and only if it can be written as

$$\rho=\sum_j p_j\rho_j^{(A)}\otimes\rho_j^{(B)} \tag{5.12}$$

Classical correlations such as the black and white marbles above fall into the category of separable states.

So far, we have considered the quantum states in the basis $\{|0\rangle,|1\rangle\}$. However, we can also describe the same system in the rotated basis $\{|\pm\rangle\}$ according to

$$|0\rangle=\frac{|+\rangle+|-\rangle}{\sqrt{2}}\quad\text{and}\quad|1\rangle=\frac{|+\rangle-|-\rangle}{\sqrt{2}} \tag{5.13}$$

The entangled state $|\Psi\rangle_{AB}$ can then be written as

$$\frac{|0,1\rangle+|1,0\rangle}{\sqrt{2}}=\frac{|+,+\rangle-|-,-\rangle}{\sqrt{2}}, \tag{5.14}$$

which means that we have again perfect correlations between the two systems with respect to the states $|+\rangle$ and $|-\rangle$. Let's do the same for the state ρ in Eq. (5.5) for classically correlated marbles. After a bit of algebra, we find that

$$\begin{aligned}\rho=\ &\frac{1}{4}(|++\rangle\langle++|+|+-\rangle\langle+-|+|-+\rangle\langle-+|+|--\rangle\langle--|\\&-|++\rangle\langle--|-|--\rangle\langle++|-|+-\rangle\langle-+|-|-+\rangle\langle+-|).\end{aligned} \tag{5.15}$$

Now there are no correlations in the conjugate basis $\{|\pm\rangle\}$, which you can check by calculating the conditional probabilities of Bob's state given Alice's measurement outcomes. This is another key difference between classically correlated states and entangled states. A good interpretation of entanglement is that entangled systems exhibit correlations that are stronger than classical correlations. We will shortly see how these stronger correlations can be used in information processing.

We have seen that operators, just like states, can be combined into tensor products:

$$A\otimes B|\phi\rangle\otimes|\psi\rangle=A|\phi\rangle\otimes B|\psi\rangle. \tag{5.16}$$

And just like states, some operators cannot be written as $A\otimes B$:

$$C=\sum_k A_k\otimes B_k \tag{5.17}$$

This is the most general expression of an operator in the Hilbert space $\mathscr{H}_1\otimes\mathscr{H}_2$. In Dirac notation this becomes

$$C=\sum_{jklm}\phi_{jklm}|\phi_j\rangle\langle\phi_k|\otimes|\phi_l\rangle\langle\phi_m|=\sum_{jklm}\phi_{jklm}|\phi_j,\phi_l\rangle\langle\phi_k,\phi_m|. \tag{5.18}$$

As an example, the Bell operator is diagonal on the Bell basis:

$$\left|\Phi^{\pm}\right\rangle=\frac{|0,0\rangle\pm|1,1\rangle}{\sqrt{2}}\text{ and }\left|\Psi^{\pm}\right\rangle=\frac{|0,1\rangle\pm|1,0\rangle}{\sqrt{2}}. \tag{5.19}$$

The eigenvalues of the Bell operator are not important, as long as they are not degenerate (why?). A measurement of the Bell operator projects onto an eigenstate of the operator, which is an entangled state. Consequently, we cannot implement such composite measurements by measuring each subsystem individually, because those individual measurements would project onto pure states of the subsystems. And we have seen that the subsystems of pure entangled states are mixed states.

A particularly useful technique when dealing with two systems is the so-called Schmidt decomposition. In general, we can write any pure state over two systems as a superposition of basis states:

$$|\Psi\rangle=\sum_{j=1}^{d_A}\sum_{k=1}^{d_B}c_{jk}|\phi_j\rangle_A|\psi_k\rangle_B, \tag{5.20}$$

where d_A and d_B are the dimensions of the Hilbert spaces of system A and B, respectively, and we order the systems such that $d_B \geq d_A$. It turns out that we can always simplify this description and write $|\Psi\rangle$ as a single sum over basis states. We state it as a theorem:

Theorem 5.2.1

Let $|\Psi\rangle$ be a pure state of two systems, A and B with Hilbert spaces $\mathscr{H}_A$ and $\mathscr{H}_B$ of dimension d_A and $d_B \geq d_A$, respectively. There exist orthonormal basis vectors $|a_j\rangle_A$ for system A and $|b_j\rangle_B$ for system B such that

$$|\Psi\rangle = \sum_j \lambda_j |a_j\rangle_A |b_j\rangle_B, \tag{5.21}$$

with real, positive Schmidt coefficients λ_j, and $\sum_j \lambda_j^2 = 1$. This decomposition is unique, and the sum runs at most to d_A, the dimension of the smallest Hilbert space. Traditionally, we order the Schmidt coefficients in descending order: $\lambda_1 \geq \lambda_2 \geq \ldots$ The total number of non-zero λ_i is the Schmidt number.

Proof

The proof can be found in many graduate texts on quantum mechanics and quantum information theory.

Given the Schmidt decomposition for a bi-partite system, we can immediately write down the reduced density matrices for the sub-systems:

$$\rho_A = \mathrm{Tr}_B(|\Psi\rangle\langle\Psi|) = \sum_j \lambda_j^2 |a_j\rangle_A \langle a_j|, \tag{5.22}$$

and

$$\rho_B = \mathrm{Tr}_A(|\Psi\rangle\langle\Psi|) = \sum_j \lambda_j^2 |b_j\rangle_B \langle b_j|. \tag{5.23}$$

The basis states $|a_j\rangle_A$ and $|b_j\rangle_B$ may have completely different physical meanings; here we care only that the states of the decomposition can be labelled with a single index, as opposed to two indices.

Conversely, when we have a single system in a mixed state

$$\rho = \sum_j p_j |a_j\rangle \langle a_j|, \tag{5.24}$$

we can always construct a pure state $|\Psi\rangle$ that obeys $\left(\lambda_j = \sqrt{p_j}\right)$

$$|\Psi\rangle = \sum_j \lambda_j |a_j, b_j\rangle, \tag{5.25}$$

By virtue of the Schmidt decomposition. The state $|\Psi\rangle$ is called the purification of ρ. Since many theorems are easier to prove for pure states than for mixed states, purifications can make our work load significantly lighter.

When there is more than one non-zero λ_j in Eq. (5.25), the state $|\Psi\rangle$ is clearly entangled: there is no alternative choice of λ_j due to the uniqueness of the Schmidt decomposition that would result in $\lambda_1' = 1$ and all others zero. Moreover, the more uniform the values of λ_j, the more the state is entangled. One possible measure for the amount of entanglement in $|\Psi\rangle$ is the Shannon entropy H.

$$H = -\sum_j \lambda_j^2 \log_2 \lambda_j^2 \tag{5.26}$$

This is identical to the von Neumann entropy S of the reduced density matrix ρ of $|\Psi\rangle$ given in Eq. (5.24):

$$S(\rho) = -\mathrm{Tr}(\rho \log_2 \rho) \tag{5.27}$$

Both entropies are measured in classical bits.

How do we find the Schmidt decomposition? Consider the state $|\Psi\rangle$ from Eq. (5.20). The (not necessarily square) matrix C with elements c_{jk} needs to be transformed into a single array of numbers λ_j. This is achieved by applying the singular-

value decomposition:

$$c_{jk} = \sum_i u_{ji} d_{ii} v_{ik}, \tag{5.28}$$

where u_{ji} and v_{ik} are elements of unitary matrices U and V, respectively, and d_{ii} is a diagonal matrix with singular values λ_i. The vectors in the Schmidt decomposition become

$$|a_i\rangle = \sum_j u_{ji} |\phi_j\rangle \quad \text{and} \quad |b_i\rangle = \sum_k v_{ik} |\psi_k\rangle . \tag{5.29}$$

This is probably a good time to remind ourselves about the singular-value decomposition. All we need to do is find U and V, the rest is just matrix multiplication. To find U, we diagonalize $CC^\dagger$ and find its eigenvectors. These form the columns of U. Similarly, we diagonalize $C^\dagger C$ and arrange the eigenvectors in columns to find V. If C is an $n \times m$ matrix, U should be $n \times n$ and V should be $m \times m$.

This page titled 5.2: Entanglement is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

5.3: Quantum teleportation

Probably the most extraordinary use of the quantum correlations present in entanglement is quantum teleportation. Alice and Bob share two entangled qubits, labelled 2 (held by Alice) and 3 (held by Bob), in the state $\left(|0,0\rangle_{23}+|1,1\rangle_{23}\right)/\sqrt{2}$. In addition, Alice holds a qubit in the

$$state|\psi\rangle_1=\alpha|0\rangle_1+\beta|1\rangle_1. \tag{5.30}$$

The object of quantum teleportation is to transfer the state of qubit 1 to qubit 3, without either Alice or Bob gaining any information about α or β. To make things extra hard, the three qubits must not change places (so Alice cannot take qubit 1 and bring it to Bob).

Classically, this is an impossible task: we cannot extract enough information about α and β with a single measurement to reproduce $|\psi\rangle$ faithfully, otherwise we could violate the no-cloning theorem. However, in quantum mechanics it can be done (without violating no-cloning). Write the total state as

$$|\chi\rangle=|\psi\rangle_1\left|\Phi^+\right\rangle_{23}=\frac{1}{\sqrt{2}}(\alpha|000\rangle+\alpha|011\rangle+\beta|100\rangle+\beta|111\rangle) \tag{5.31}$$

Alice now performs a Bell measurement on her two qubits (1 and 2), which project them onto one of the Bell states $\left|\Phi^\pm\right\rangle_{12}$ or $\left|\Psi^\pm\right\rangle_{12}$. We write $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$ in the Bell basis:

$$\begin{aligned}
|00\rangle&=\frac{|\Phi^+\rangle+|\Phi^-\rangle}{\sqrt{2}}\\
|01\rangle&=\frac{|\Psi^+\rangle+|\Psi^-\rangle}{\sqrt{2}}\\
|10\rangle&=\frac{|\Psi^+\rangle-|\Psi^-\rangle}{\sqrt{2}}\\
|11\rangle&=\frac{|\Phi^+\rangle-|\Phi^-\rangle}{\sqrt{2}}.
\end{aligned} \tag{5.32}$$

We can use these substitutions to write the state $|\chi\rangle$ before the measurement as

$$\begin{aligned}
|\chi\rangle=&\frac{1}{2}\left[\alpha\left|\Phi^+\right\rangle_{12}|0\rangle_3+\alpha\left|\Phi^-\right\rangle_{12}|0\rangle_3+\alpha\left|\Psi^+\right\rangle_{12}|1\rangle_3+\alpha\left|\Psi^-\right\rangle_{12}|1\rangle_3\right.\\
&\left.+\beta\left|\Psi^+\right\rangle_{12}|0\rangle_3-\beta\left|\Psi^-\right\rangle_{12}|0\rangle_3+\beta\left|\Phi^+\right\rangle_{12}|1\rangle_3-\beta\left|\Phi^-\right\rangle_{12}|1\rangle_3\right]\\
=&\frac{1}{2}\left[\left|\Phi^+\right\rangle(\alpha|0\rangle+\beta|1\rangle)+\left|\Phi^-\right\rangle(\alpha|0\rangle-\beta|1\rangle)\right.\\
&\left.+\left|\Psi^+\right\rangle(\beta|0\rangle+\alpha|1\rangle)+\left|\Psi^-\right\rangle(\beta|0\rangle-\alpha|1\rangle)\right]
\end{aligned} \tag{5.33}$$

Alice finds one of four possible outcomes:

$$\begin{aligned}
\Phi^+:&\quad \mathrm{Tr}_{12}(|\Phi^+\rangle\langle\Phi^+||\chi\rangle\langle\chi|)\rightarrow|\psi\rangle_3=\alpha|0\rangle+\beta|1\rangle\\
\Phi^-:&\quad \mathrm{Tr}_{12}(|\Phi^-\rangle\langle\Phi^-||\chi\rangle\langle\chi|)\rightarrow|\psi\rangle_3=\alpha|0\rangle-\beta|1\rangle\\
\Psi^+:&\quad \mathrm{Tr}_{12}(|\Psi^+\rangle\langle\Psi^+||\chi\rangle\langle\chi|)\rightarrow|\psi\rangle_3=\alpha|1\rangle+\beta|0\rangle\\
\Psi^-:&\quad \mathrm{Tr}_{12}(|\Psi^-\rangle\langle\Psi^-||\chi\rangle\langle\chi|)\rightarrow|\psi\rangle_3=\alpha|1\rangle-\beta|0\rangle
\end{aligned} \tag{5.34}$$

From these outcomes, it is clear that the state held by Bob is different for the different measurement outcomes of Alice's Bell measurement. Let this sink in for a moment: After setting up the entangled state between Alice and Bob, who may be literally light years apart, Bob has done absolutely nothing to his qubit, yet its state is different depending on Alice's measurement outcome. This suggests that there is some instantaneous communication taking place, possibly violating causality!

In order to turn the state of Bob's qubit into the original state, Alice needs to send the measurement outcome to Bob. This will take two classical bits, because there are four outcomes. The correction operators that Bob need to apply are as follows:

$$\Phi^+:\mathbb{I},\quad \Phi^-:Z,\quad \Psi^+:X,\quad \Psi^-:ZX \tag{5.35}$$

So in each case Bob needs to do something different to his qubit. To appreciate how remarkable this protocol is, here are some of its relevant properties:

1. No matter is transported, only the state of the system;
2. neither Alice nor Bob learns anything about α or β;
3. any attempt to use quantum teleportation for signaling faster than light is futile!

? Exercises

1. Derive Eq. (5.15), and show that $\rho_B = \mathbb{I}/2$ when Alice's qubit is projected onto $|+\rangle$.
2. Quantum teleportation. Write the Bell states as

$$|\psi_{nm}\rangle = \left(|0, 0 \oplus n\rangle + (-1)^m |1, 1 \oplus n\rangle\right)/\sqrt{2},$$

   where $\oplus$ denotes addition modulo 2 and $n, m = 0, 1$.

   a. Write $|\psi_{nm}\rangle$ in terms of $|\psi_{00}\rangle$ and the Pauli operators X and Z acting on the second qubit.
   b. Using the shared Bell state $|\psi_{nm}\rangle$ between Alice and Bob, and the two-bit measurement outcome (j, k) for Alice's Bell measurement, determine the correction operator for Bob.
   c. We now generalize to N-dimensional systems. We define the N^2 entangled states

$$|\psi_{nm}\rangle = \tfrac{1}{\sqrt{N}} \textstyle\sum_{j=0}^{N-1} e^{2\pi i j n/N} |j, j \oplus m\rangle \ ,$$

      where $n \oplus m = n + m \mod N$. Prove that this is an orthonormal basis.

   d. Give the teleportation protocol for the N-dimensional Hilbert space.
   e. What is Bob's state before he learns Alice's measurement outcome?

3. Imperfect measurements.

   a. A two-qubit system (held by Alice and Bob) is in the anti-symmetric Bell state $|\Psi^-\rangle$. Calculate the state of Bob's qubit if Alice measures her qubit in the state $|0\rangle$. Hint: write the measurement procedure as a partial trace over Alice's qubit.
   b. Now Alice's measurement is imperfect, and when her apparatus indicates "0", there was actually a small probability p that the qubit was projected onto the state $|1\rangle$. What is Bob's state?
   c. If Alice's (imperfect) apparatus has only two measurement outcomes, what will Bob's state be if she finds outcome "1"?

This page titled 5.3: Quantum teleportation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

6: Evolution of Open Quantum Systems

We have considered mixed states, where the experimenter has incomplete information about the state preparation procedure, and we have also seen that mixing arises in a system when it is entangled with another system. The combined system can still be pure, but the subsystem has become mixed. This phenomenon arises often when we want to describe systems that have some interaction with their environment. The interaction creates entanglement, and the system taken by itself evolves from a pure state to a mixed state. Such a system is called “open”, since it can leak quantum information to the environment. The theory of open quantum systems revolves around the so-called Lindblad equation.

6.1: The Lindblad Equation

6.2: Positive and Completely Positive Maps

6.3: Bra Vectors and the Inner Product

6.4: Normalization and Orthogonality

6.5: Interpreting the Inner Product

6.5.1: Propagating Amplitudes

This page titled 6: Evolution of Open Quantum Systems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

6.1: The Lindblad Equation

Next, we will derive the Lindblad equation, which is the direct extension of the Heisenberg equation for the density operator, i.e., the mixed state of a system. We have seen in Eq. (4.30) that formally, we can write the evolution of a density operator as a mathematical map $\mathscr{E}$, such that the density operator ρ transforms into

$$\rho \to \rho' = \mathscr{E}(\rho) \equiv \sum_k A_k \rho A_k^\dagger, \tag{6.1}$$

where the A_k are the Kraus operators. Requiring that ρ' is again a density operator $(\mathrm{Tr}(\rho') = 1)$ leads to the restriction $\sum_k A_k^\dagger A_k = \mathbb{I}$.

We want to describe an infinitesimal evolution of ρ, in order to give the continuum evolution later on. We therefore have that

$$\rho' = \rho + \delta\rho = \sum_k A_k \rho A_k^\dagger \tag{6.2}$$

Since $\delta\rho$ is very small, one of the Kraus operators must be close to the identity. Without loss of generality we choose this to be A_0, and then we can write

$$A_0 = \mathbb{I} + (L_0 - iK)\,\delta t \quad \text{and} \quad A_k = L_k\sqrt{\delta t}, \tag{6.3}$$

where we introduced the Hermitian operators L_0 and K, and the remaining L_k are not necessarily Hermitian. We could have written $A_0 = \mathbb{I} + L_0\delta t$ and keep L_0 general (non-Hermitian as well), but it will be useful later on to explicitly decompose it into Hermitian parts. We can now write

$$\begin{aligned} A_0 \rho A_0^\dagger &= \rho + \left[(L_0 - iK)\,\rho + \rho\,(L_0 + iK)\right]\delta t + O\left(\delta t^2\right) \\ A_k \rho A_k^\dagger &= L_k \rho L_k^\dagger \delta t \end{aligned} \quad . \tag{6.4}$$

We can substitute this into Eq. (6.2), to obtain up to first order in δt

$$\delta\rho = \left[(L_0\rho + \rho L_0) - i(K\rho - \rho K) + \sum_{k\neq 0} L_k \rho L_k^\dagger\right]\delta t. \tag{6.5}$$

We now give the continuum evolution by dividing by δt and taking the limit $\delta t \to dt$:

$$\frac{d\rho}{dt} = -i[K,\rho] + \{L_0,\rho\} + \sum_{k\neq 0} L_k \rho L_k^\dagger, \tag{6.6}$$

where $\{A,B\} = AB + BA$ is the anti-commutator of A and B. We are almost there, but we must determine what the different terms mean. Suppose we consider the free evolution of the system.

Eq. (6.6) must then reduce to the Heisenberg equation for the density operator ρ in Eq. (4.7), and we see that all L_k including L_0 are zero, and K is proportional to the Hamiltonian $K = H/\hbar$. Again from the general property that $\mathrm{Tr}(\rho) = 1$ we have

$$\mathrm{Tr}\left(\frac{d\rho}{dt}\right) = 0 \to L_0 = -\frac{1}{2}\sum_{k\neq 0} L_k^\dagger L_k. \tag{6.7}$$

This finally leads to the Lindblad equation

$$\frac{d\rho}{dt} = \frac{1}{i\hbar}[H,\rho] + \frac{1}{2}\sum_k \left(2L_k\rho L_k^\dagger - \left\{L_k^\dagger L_k, \rho\right\}\right). \tag{6.8}$$

The operators L_k are chosen such that they model the relevant physical processes. This may sound vague, but in practice it will be quite clear. For example, modelling a transition $|1\rangle \to |0\rangle$ without keeping track of where the energy is going or coming from will require a single Lindblad operator

$$L = \gamma|0\rangle\langle 1|, \tag{6.9}$$

where γ is a real parameter indicating the strength of the transition. This can model both decay and excitations.

This page titled 6.1: The Lindblad Equation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

6.2: Positive and Completely Positive Maps

We considered the evolution of the density operator under a family of Kraus operators in Eq. (4.30):

$$\rho \to \rho' = \mathscr{E}(\rho) = \sum_k A_k \rho A_k^\dagger, \tag{6.10}$$

where $\sum_k A_k^\dagger A_k = \mathbb{I}$ (that is, $\mathscr{E}$ is trace-preserving). When $\mathscr{E}$ transforms any positive operator into another positive operator, we call it a positive map. We may be tempted to conclude that all positive maps correspond to physically allowed transformations. After all, it maps density operators to density operators. Unfortunately, Nature (or Mathematics?) is not that tidy.

Consider the transpose of the density operator $\rho \to \rho^T$, which acts according to

$$\rho = \sum_{ij} \rho_{ij} |i\rangle\langle j| \to \rho^T = \sum_{ij} \rho_{ji} |i\rangle\langle j|. \tag{6.11}$$

You can verify immediately that the trace is preserved in this operation (check this!), and ρ^T is again a positive operator since the eigenvalues are identical to those of ρ. For example, consider the qubit state $(|0\rangle + i|1\rangle)/\sqrt{2}$. The density operator and its transpose are

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \quad \text{and} \quad \rho^T = \frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}. \tag{6.12}$$

The transpose therefore corresponds to the state $(|0\rangle - i|1\rangle)/\sqrt{2}$. Now consider that the qubit is part of an entangled state $(|00\rangle + |11\rangle)/\sqrt{2}$. The density operator is given by

$$\rho = \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|), \tag{6.13}$$

and the partial transpose on the first qubit is

$$\rho^T = \frac{1}{2}(|00\rangle\langle 00| + |10\rangle\langle 01| + |01\rangle\langle 10| + |11\rangle\langle 11|). \tag{6.14}$$

The eigenvalues of ρ are all positive, but ρ^T has a negative eigenvalue! So ρ^T cannot be a density operator. Consequently, it is not correct to say that positive maps correspond to physical processes. We need to put another restriction on maps.

From the example of the partial transpose, we can deduce that maps must not only be positive for the system S that they act on, but also positive on larger systems that include S as a subsystem. When this is the case, we call the map completely positive. There is a very important theorem in mathematics, called Kraus' Representation Theorem, which states that maps of the form in Eq. (6.10) with the restriction that $\sum_k A_k^\dagger A_k = \mathbb{I}$ is a completely positive map, and moreover, that any completely positive map can be expressed in this form.

? Exercises

1. a. Show that $\sum_k A_k^\dagger A_k = \mathbb{I}$,
   b. prove that any non-Hermitian square matrix can be written as $A + iB$, with A and B Hermitian,
   c. prove that $L_0 = -\frac{1}{2}\sum_{k\neq 0} L_k^\dagger L_k$.
2. Consider a two-level system $(|0\rangle, |1\rangle)$ that has a dephasing process, modelled by the Lindblad operators $L_1 = \gamma|0\rangle\langle 1|$ and $L_2 = \gamma|1\rangle\langle 0|$.
   a. write down the Lindblad equation (choose $H = 0$ for simplicity).
   b. Calculate the evolution of the pure states $|0\rangle$ and $|+\rangle$ at $t = 0$. Hint: write the density matrix in the Pauli matrix basis $\{\mathbb{I}, X, Y, Z\}$. What can you say about the equilibrium state of the system?
   c. Calculate and plot the entropy $S(\rho)$ of the state $\rho(t)$ as a function of γ and t.
3. Calculate the eigenvalues of ρ^T in Eq. (6.14).

This page titled 6.2: Positive and Completely Positive Maps is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available

upon request.

6.3: Bra Vectors and the Inner Product

For each ket vector $|\psi\rangle$, there is a corresponding bra vector $\langle\psi|$. We haven't yet looked into any specific representations of ket vectors beyond just the ket vector itself, so at the moment that's all you need to know. However, when we do get into specific representations, the rules for converting ket vectors to bra vectors are generally very easy. You always take the complex conjugate of any numbers in the representation going from the ket vector to the bra vector. (You may also turn a column vector into a row vector, if you're using column vectors to represent ket vectors; much more about that later.) $\langle\psi|$ hψ| is something like the complex conjugat $|\psi\rangle$, although that's not really right. However, just as a number and its complex conjugate are associated with each other, each ket vector $|\psi\rangle$ is uniquely associated with a bra vetor $\langle\psi|$.

With the introduction of bra vectors, it becomes possible to define a new operation you can do on these things. You can always stick a bra vector on to a ket vector. The notation is meant to help suggest this; where there is a straight side, you can stick two of them together. The result is called the inner product. The specific rules for how you calculate the inner product again depend on the detailed representation of the ket vector, so for now we'll keep them abstract. As an example, suppose you have two different quantum states represented by the ket vector $|\psi\rangle$ and the ket vector $|\phi\rangle$. The bra vector corresponding to the latter is $\langle\phi|$, and the inner product of that bra vector with the ket vector $|\psi\rangle$ is:

$$\langle\phi\,|\,\psi\rangle$$

When you see a bra-ket pair combined like that, the result is a **scalar!** It may well be a complex number, but it is just a number. At that point, you can manipulate it in algebraic equations the way you would manipulate any other complex number.

The inner product of a bra and a ket is the first way we've seen to multiply two of these state vectors together. We've talked about multiplying the state vectors by a scalar, but before we didn't know how to multiply them together. Notice, however, that this is a different sort of multiplication than multiplying two scalars. When you multiply two scalars, you get another scalar out— the same sort of thing as the things you multiplied together. However, when you take the inner product of two state vectors, you get a scalar out, something different from the two things that went into the inner product.

Note that you can only take the inner product between two quantum states if they are the same sort of state. That is, they must be the same kind of state for the same particle or system. For instance, you could take the inner product between two angular momentum states for the same electron, but you couldn't take the inner product between an angular momentum state and a position state.

This page titled 6.3: Bra Vectors and the Inner Product is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

6.4: Normalization and Orthogonality

Although we aren't yet going to learn rules for doing general inner products between state vectors, there are two cases where the inner product of two state vectors produces a simple answer. The first is not intrinsic to the mathematical representation, but rather something we will insist for state vectors that properly represent real physical states. For a complete state vector $|\psi\rangle$ to be a proper quantum mechanical state, it must satisfy the condition

$$\langle \psi \mid \psi \rangle = 1$$

We say that this means that the state vector is normalized. It is possible to have non-normalized state vectors. For instance, in the equation

$$|+x\rangle = \tfrac{1}{\sqrt{2}}|+z\rangle + \tfrac{1}{\sqrt{2}}|-z\rangle$$

the two parts of the sum on the right side are themselves ket vectors. However, because they are valid state vectors multiplied by a constant, they are not normalized themselves. We will show later that this definition of $|+x\rangle$ is, however, normalized.

The second rule is that state vectors that represent different possible states corresponding to different possible measurements of a given observable must be orthogonal. Mathematically, this is expressed as:

$$\langle \phi_1 \mid \phi_2 \rangle = 0$$

if $|\phi_1\rangle$ and $|\phi_2\rangle$ are two different states corresponding to definite states for a given observable. For example, the states $|+z\rangle$ and $|-z\rangle$ correspond to two states of the same observable, specifically, the z component of angular momentum. The first corresponds to that component being measured along $+z$, the second to it being measured along $-z$. The orthogonality condition is then:

$$\langle +z \mid -z \rangle = 0$$

As an example of doing these calculations with a more complicated state, consider the state $|+x\rangle$. If this state is properly normalized, then we should have $\langle +x \mid +x \rangle = 1$. Do we? Well, first, we have to construct the bra vector that goes along with the ket vector:

$$|+x\rangle = \frac{1}{\sqrt{2}}|+z\rangle + \frac{1}{\sqrt{2}}|-z\rangle$$
$$\langle +x| = \frac{1}{\sqrt{2}}|+z\rangle + \frac{1}{\sqrt{2}}|-z\rangle$$

Notice that in the case of a compound ket vector, to get the bra vector we just turn all ket vectors on the right side into bra vectors, and replace all the numbers with their complex conjugates (which is trivial here, since all the numbers are real). Now we have what we need to figure out the inner product. Just substitute in our expressions for $|+x\rangle$ and $\langle +x|$:

$$\begin{aligned}\langle +x \mid +x \rangle &= \left(\frac{1}{\sqrt{2}}\left\langle +z\right| + \frac{1}{\sqrt{2}}\langle -z|\right)\left(\frac{1}{\sqrt{2}}|+z\rangle + \frac{1}{\sqrt{2}}|-z\rangle\right) \\ &= \frac{1}{2}\langle +z \mid +z\rangle + \frac{1}{2}\langle +z \mid -z\rangle + \frac{1}{2}\langle -z \mid +z\rangle + \frac{1}{2}\langle -z \mid -z\rangle\end{aligned}$$

That looks very complicated, but now we can use the orthogonality condition we know is true for the z states, as we've defined them as good states corresponding to the z component of z angular momentum. We know that $\langle +z \mid +z\rangle = 1, \langle -z \mid -z\rangle = 1$, $\langle -z \mid +z\rangle = 0$, and $\langle +z \mid -z\rangle = 0$ from normalization and orthogonality. Substitute these in:

$$\begin{aligned}\langle +x \mid +x \rangle &= \frac{1}{2}(1) + \frac{1}{2}(0) + \frac{1}{2}(0) + \frac{1}{2}(1) \\ \langle +x \mid +x \rangle &= 1\end{aligned}$$

So the state is properly normalized! I leave it as an exercise for the alert reader to show that $|+x\rangle$ and $|-x\rangle$ are orthogonal.

This page titled 6.4: Normalization and Orthogonality is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

6.5: Interpreting the Inner Product

So far, all we know about the inner product is that for a properly normalized quantum state, the inner product of that state with itself is 1, and that the inner product between two different states corresponding to definite states of the same observable must be zero. But what about the inner product between two arbitrary states? Consider:

$$\langle \phi \,|\, \psi \rangle$$

The interpretation of this is that it is the **amplitude for a particle in state $|\psi\rangle$ to subsequently be measured in state $|\phi\rangle$**. As an example, suppose that we have an electron in the following state:

$$|\psi\rangle = \tfrac{3}{5}|+z\rangle + \tfrac{4i}{5}|-z\rangle$$

Suppose we send this electron through an SGz machine. If this state is properly normalized (is it?), then we could work out the amplitude for it to be measured in the $|-z\rangle$ state (i.e. the amplitude for measuring its z-spin to be $-\hbar/2$) as follows:

$$\begin{aligned}
\langle -z \,|\, \psi \rangle &= \langle -z| \left(\frac{3}{5}|+z\rangle + \frac{4i}{5}|-z\rangle \right) \\
&= \frac{3}{5}\langle -z \,|\, +z\rangle + \frac{4i}{5}\langle -z \,|\, -z\rangle \\
&= \frac{3}{5}(0) + \frac{4i}{5}(1) \\
&= \frac{4i}{5}
\end{aligned}$$

This tells us the amplitude for the electron to be found in the $|-z\rangle$ state. Remember that the probability, what we really care about, is the absolute square of the amplitude. That probability is:

$$\begin{aligned}
|\langle -z \,|\, \psi \rangle|^2 &= \langle -z \,|\, \psi \rangle^* \langle -z \,|\, \psi \rangle \\
&= \left(\frac{-4i}{5} \right) \left(\frac{4i}{5} \right) \\
&= \left(\frac{-16i^2}{25} \right) \\
&= \frac{16}{25} = 0.64
\end{aligned}$$

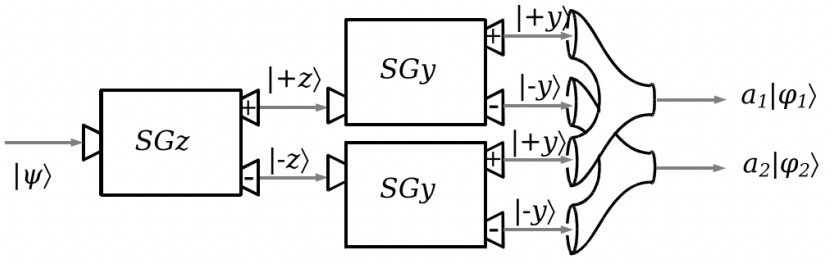
If the electron was in the state $|\psi\rangle$ defined above upon entering a SGz machine, there's an 64% chance it will come out the $-z$ output of the machine, being measured with a z-spin of $-\hbar/2$.

This page titled 6.5: Interpreting the Inner Product is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

6.5.1: Propagating Amplitudes

We have seen that the amplitude for a given quantum state $|\psi\rangle$ to later be found in another quantum state $|\phi\rangle$ is $\langle\phi \mid \psi\rangle$. Physically, when would you apply this amplitude? You would apply it when the system went through a device that measured whatever quantity $|\phi\rangle$ is associated with. For example, if you have an electron beam in state $|\psi\rangle$ going into an SGz machine, you'd associate the amplitude $\langle +z \mid \psi\rangle$ with the state $|+z\rangle$ emerging from the positive output of the machine. What do you do, however, if the electron beam then goes through another machine? How do you deal with amplitudes when there is more than one process that might have a state change associated with it? The answer is that to get the overall amplitude for a starting state to end up in some final state, you multiply the individual amplitudes of each step the system went through.[2]

As an example, consider the following sequence of SG machines:



Yowza.[3] An electron in some state $|\psi\rangle$ goes into the beginning of this system. There are two possible places it may come out. It may emerge from the upper output in state $|\phi_1\rangle$ (which is currently unknown, but we will figure it out); the amplitude for it to emerge here is a_1. It may also emerge from the lower output in state $|\phi_2\rangle$ (which we will also figure out); the amplitude for it to emerge from the lower output is a_2.

Ultimately, what we're interested in is the amplitude for the electron emerging from this whole thing with state $|+y\rangle$, and the amplitude for it emerging with state $|-y\rangle$. To figure those out, we need to trace the electron through all of the possible paths. At the input to the first machine, the electron is in the state $|\psi\rangle$. At the positive output of the first machine, the electron is now either in the state $|+z\rangle$, with amplitude $\langle +z \mid \psi\rangle$, or in the state $|-z\rangle$, with amplitude $\langle -z \mid \psi\rangle$.

Let's consider the possible paths for the electron if it comes out of the $+z$ output of the first machine. If the electron goes this way, it will go into the upper SGy machine, with state $|+z\rangle$. It will emerge from either the + output, with amplitude $\langle +y \mid +z\rangle$, or from the − output, with amplitude $\langle -y \mid +z\rangle$. The overall amplitude for the electron to make it from the very beginning to the + output of the upper second machine is the product of the amplitudes for each step: $\langle +z \mid \psi\rangle\langle +y \mid +z\rangle$. Likewise, the overall amplitude for the electron to make it from the very beginning to the − output for the lower second machine is $\langle +z \mid \psi\rangle\langle +y \mid +z\rangle$.

Next, consider the possible path of the electron emerging from the − output of the first machine. The amplitude for it to get this far is $\langle -z \mid \psi\rangle$. The overall amplitude, then, for it to come out of the + output of the lower machine is $\langle -z \mid \psi\rangle\langle +y \mid -z\rangle$, and the overall amplitude for it to come out of the − output of the lower machine is $\langle -z \mid \psi\rangle\langle -y \mid -z\rangle$.

What do you do at a beam combiner? There, you just add the two states together, each multiplied by their respective amplitudes. Let's first consider the upper beam combiner. The two states coming into this system, with their respective amplitudes, are:

$$\langle +z \mid \psi\rangle\langle +y \mid +z\rangle|+y\rangle$$

and

$$\langle -z \mid \psi\rangle\langle +y \mid -z\rangle|+y\rangle \ .$$

Therefore, the final output amplitude and state is:

$$a_1\,|\phi_1\rangle = [\langle +z \mid \psi\rangle\langle +y \mid +z\rangle + \langle -z \mid \psi\rangle\langle +y \mid -z\rangle]|+y\rangle$$

By looking at this, we can see that the state $|\phi_1\rangle$ is in fact just $|+y\rangle$. Hopefully, that does not come as a surprise to you, as the state of the two electron beams going into this beam combiner was just $|+y\rangle$. The amplitude a_1 is then just

$$a_1 = \langle +z \mid \psi\rangle\langle +y \mid +z\rangle + \langle -z \mid \psi\rangle\langle +y \mid -z\rangle$$

The two state going into the lower beam combiner, with their respective amplitudes, are:

$$\langle +z \mid \psi\rangle\langle -y \mid +z\rangle|-y\rangle$$

and

$$\langle -z \mid \psi \rangle \langle -y \mid -z \rangle | -y \rangle \ .$$

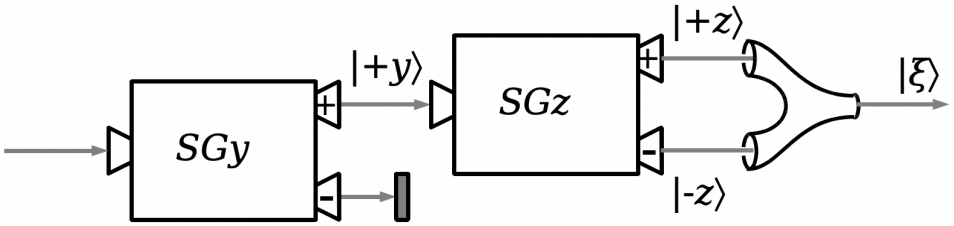
Therefore, the final output amplitude and state for the lower output from this whole system is:

$$a_2 \, |\phi_2\rangle = [\langle +z \mid \psi \rangle \langle -y \mid +z \rangle + \langle -z \mid \psi \rangle \langle -y \mid -z \rangle] | -y \rangle$$

The state $|\phi_2\rangle$, and the amplitude for the lower output is:

$$a_2 = \langle +z \mid \psi \rangle \langle -y \mid +z \rangle + \langle -z \mid \psi \rangle \langle -y \mid -z \rangle$$

As another example, consider the following collection of SG machines:



We know from the previous chapter that the final state of this system should be $|+y\rangle$. Is that what we get?

In order to analyze this, you're going to need to know how to express the states $|+y\rangle$ and $|-y\rangle$ in terms of $|+z\rangle$ and $|-z\rangle$:

$$|+y\rangle = \frac{1}{\sqrt{2}}|+z\rangle + \frac{i}{\sqrt{2}}|-z\rangle$$
$$|-y\rangle = \frac{i}{\sqrt{2}}|+z\rangle + \frac{1}{\sqrt{2}}|-z\rangle$$

Consider the electron going into the input of the second machine. It is in state $|+y\rangle$. We won't worry about the amplitude for the initial electron to get into this state, because we'll just consider the ones that happen to come out the positive output of the first machine. (The purpose of that first machine is to make sure that we know the electrons are in fact in the $|+y\rangle$ state when they go into the second machine.) The amplitude for an electron to come out of the upper terminal is $\langle +z \mid +y \rangle$, and the amplitude for an electron to come out of the lower terminal is $\langle -z \mid +y \rangle$. Call the final state coming out of the beam combiner $|\xi\rangle$. To figure out what this state is, combine together the two states going into it, each multiplied by their respective amplitudes:

$$\begin{aligned}
|\xi\rangle &= \langle +z \mid +y \rangle | +z \rangle + \langle -z \mid +y \rangle | -z \rangle \\
&= \langle +z | \left(\frac{1}{\sqrt{2}}|+z\rangle + \frac{i}{\sqrt{2}}|-z\rangle \right) |+z\rangle + \langle -z | \left(\frac{1}{\sqrt{2}}|+z\rangle + \frac{i}{\sqrt{2}}|-z\rangle \right) |-z\rangle \\
&= \left(\frac{1}{\sqrt{2}}\langle +z \mid +z \rangle + \frac{i}{\sqrt{2}}\langle +z \mid -z \rangle \right) |+z\rangle + \left(\frac{1}{\sqrt{2}}\langle -z \mid +z \rangle + \frac{i}{\sqrt{2}}\langle -z \mid -z \rangle \right) |-z\rangle
\end{aligned}$$

Once again, we just have inner products of z states with themselves. We can use normalization (e.g. $\langle +z \mid +z \rangle = 1$) and orthogonality (e.g. $\langle -z \mid +z \rangle = 0$) to substitute in the numbers from the inner products in the expression above, yielding us:

$$\begin{aligned}
|\xi\rangle &= \left(\frac{1}{\sqrt{2}}(1) + \frac{i}{\sqrt{2}}(0) \right) |+z\rangle + \left(\frac{1}{\sqrt{2}}(0) + \frac{i}{\sqrt{2}}(1) \right) |-z\rangle \\
&= \frac{1}{\sqrt{2}}|+z\rangle + \frac{i}{\sqrt{2}}|-z\rangle \\
&= |+y\rangle
\end{aligned}$$

Sure enough, the mathematical rules for propagating amplitudes has given us the answer that we know is supposed to be right for the final state.

---

[2]This is different from classical physics, where you'd multiply the probabilities. You might wonder what the difference is, since you are going to square the whole thing at the end anyway. The differences comes from the fact that the quantum amplitudes may be complex, so the products of the individual amplitudes could end up with terms canceling each other.

[3]The reason why we have this complicated a collection of SG machines and beam combiners is that it's important that we not be able to figure out which output from the SGz machine the electron went through, for subtle reasons that will be discussed in the

next chapter.

This page titled 6.5.1: Propagating Amplitudes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

7: Orbital Angular Momentum and Spin

Angular momentum plays an important role in quantum mechanics, not only as the orbital angular momentum of electrons orbiting the central potentials of nuclei, but also as the intrinsic magnetic moment of particles, known as spin, and even as isospin in high-energy particle physics.

7.1: Orbital Angular Momentum

7.2: Spin

7.3: Total Angular Momentum

7.4: Composite Systems with Angular Momentum

This page titled 7: Orbital Angular Momentum and Spin is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

7.1: Orbital Angular Momentum

From classical physics we know that the orbital angular momentum of a particle is given by the cross product of its position and momentum

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad \text{or} \quad L_i = \epsilon_{ijk} r_j p_k, \tag{7.1}$$

where we used Einstein's summation convention for the indices. In quantum mechanics, we can find the operator for orbital angular momentum by promoting the position and momentum observables to operators. The resulting orbital angular momentum operator turns out to be rather complicated, due to a combination of the cross product and the fact that position and momentum do not commute. As a result, the components of orbital momentum do not commute with each other. When we use $[r_j, p_k] = i\hbar\delta_{jk}$, the commutation relation for the components of $\mathbf{L}$ becomes

$$[L_i, L_j] = i\hbar\epsilon_{ijk} L_k. \tag{7.2}$$

A set of relations like this is called an algebra, and the algebra here is called closed since we can take the commutator of any two elements L_i and L_j, and express it in terms of another element L_k. Another (simpler) closed algebra is $[x, p_x] = i\hbar\mathbb{I}$ and $[x, \mathbb{I}] = [p_x, \mathbb{I}] = 0$.

Since the components of angular momentum do not commute, we cannot find simultaneous eigenstates for L_x, L_y, and L_z. We will choose one of them, traditionally denoted by L_z, and construct its eigenstates. It turns out that there is another operator, a function of all L_is, that commutes with any component L_j, namely $\mathbf{L}^2 = L_x^2 + L_y^2 + L_z^2$. This operator is unique, in that there is no other operator that differs from L^2 in a nontrivial way and still commutes with all L_is. We can now construct simultaneous eigenvectors for L_z and $\mathbf{L}^2$.

Since we are looking for simultaneous eigenvectors for the square of the angular momentum and the z-component, we expect that the eigenvectors will be determined by two quantum numbers, l, and m. First, and without any prior knowledge, we can formally write down the eigenvalue equation for L_z as

$$L_z|l, m\rangle = m\hbar|l, m\rangle, \tag{7.3}$$

where m is some real number, and we included $\hbar$ to fit the dimensions of angular momentum. We will now proceed with the derivation of the eigenvalue equation for $\mathbf{L}^2$, and determine the possible values for l and m.

From the definition of L^2, we have $\mathbf{L}^2 - L_z^2 = L_x^2 + L_y^2$, and

$$\langle l, m | \mathbf{L}^2 - L_z^2 | l, m\rangle = \langle l, m | L_x^2 + L_y^2 | l, m\rangle \geq 0 \tag{7.4}$$

The spectrum of L_z is therefore bounded by

$$l \leq m \leq l \tag{7.5}$$

for some value of l. We derive the eigenvalues of $\mathbf{L}^2$ given these restrictions. First, we define the ladder operators

$$L_\pm = L_x \pm iL_y \quad \text{with} \quad L_- = L_+^\dagger. \tag{7.6}$$

The commutation relations with L_z and $\mathbf{L}^2$ are

$$[L_z, L_\pm] = \pm\hbar L_\pm, \quad [L_+, L_-] = 2\hbar L_z, \quad \left[L_\pm, \mathbf{L}^2\right] = 0. \tag{7.7}$$

Next, we calculate $L_z\left(L_+|l, m\rangle\right)$:

$$\begin{aligned} L_z\left(L_+|l, m\rangle\right) &= \left(L_+L_z + [L_z, L_+]\right)|l, m\rangle = m\hbar L_+|l, m\rangle + \hbar L_+|l, m\rangle \\ &= (m+1)\hbar L_+|l, m\rangle. \end{aligned} \tag{7.8}$$

Therefore $L_+|l, m\rangle \propto |l, m+1\rangle$. By similar reasoning we find that $L_-|l, m\rangle \propto |l, m-1\rangle$. Since we already determined that $-l \leq m \leq l$, we must also require that

$$L_+|l, l\rangle = 0 \quad \text{and} \quad L_-|l, -l\rangle = 0. \tag{7.9}$$

Counting the states between $-l$ and $+l$ in steps of one, we find that there are $2l+1$ different eigenstates for L_z. Since $2l+1$ is a positive integer, l must be a half-integer $\left(l = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots\right)$. Later we will restrict this further to $l = 0, 1, 2, \ldots$

The next step towards finding the eigenvalues of $\mathbf{L}^2$ is to calculate the following identity:

$$L_- L_+ = (L_x - iL_y)(L_x + iL_y) = L_x^2 + L_y^2 + i[L_x, L_y] = \mathbf{L}^2 - L_z^2 - \hbar L_z \tag{7.10}$$

We can then evaluate

$$L_- L_+ |l,l\rangle = 0 \Rightarrow \left(\mathbf{L}^2 - L_z^2 - \hbar L_z\right)|l,l\rangle = \mathbf{L}^2|l,l\rangle - \left(l^2 + l\right)\hbar^2 |l,l\rangle = 0 \tag{7.11}$$

It is left as an exercise (see exercise 1b) to show that

$$\mathbf{L}^2|l,m\rangle = l(l+1)\hbar^2|l,m\rangle. \tag{7.12}$$

We now have derived the eigenvalues for L_z and $\mathbf{L}^2$.

One aspect of our algebraic treatment of angular momentum we still have to determine is the matrix elements of the ladder operators. We again use the relation between $L_\pm$, and L_z and $\mathbf{L}^2$:

$$\langle l,m|L_- L_+|l,m\rangle = \sum_{j=-l}^{l} \langle l,m|L_-|l,j\rangle \langle l,j|L_+|l,m\rangle . \tag{7.13}$$

Both sides can be rewritten as

$$\left\langle l,m\left|\mathbf{L}^2 - L_z^2 - \hbar L_z\right|l,m\right\rangle = \langle l,m|L_-|l,m+1\rangle \langle l,m+1|L_+|l,m\rangle , \tag{7.14}$$

where on the right-hand-side we used that only the $m+1$-term survives. This leads to

$$[l(l+1) - m(m+1)]\hbar^2 = |\langle l,m+1|L_+|l,m\rangle|^2. \tag{7.15}$$

The ladder operators then act as

$$L_+|l,m\rangle = \hbar\sqrt{l(l+1) - m(m+1)}\,|l,m+1\rangle, \tag{7.16}$$

and

$$L_-|l,m\rangle = \hbar\sqrt{l(l+1) - m(m-1)}\,|l,m-1\rangle. \tag{7.17}$$

We have seen that the angular momentum L is quantized, and that this gives rise to a discrete state space parameterized by the quantum numbers l and m. However, we still have to restrict the values of l further, as mentioned above. We cannot do this using only the algebraic approach (i.e., using the commutation relations for L_i), and we have to consider the spatial properties of angular momentum. To this end, we write L_i as

$$L_i = -i\hbar\epsilon_{ijk}\left(x_j \frac{\partial}{\partial x_k}\right), \tag{7.18}$$

which follows directly from the promotion of $\mathbf{r}$ and $\mathbf{p}$ in Eq. (7.1) to quantum mechanical operators. In spherical coordinates,

$$r = \sqrt{x^2 + y^2 + z^2}\,, \quad \phi = \arctan\left(\frac{y}{x}\right), \quad \theta = \arctan\left(\frac{\sqrt{x^2+y^2}}{z}\right), \tag{7.19}$$

the angular momentum operators can be written as

$$\begin{aligned}
L_x &= -i\hbar\left(-\sin\phi\frac{\partial}{\partial\theta} - \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right) \\
L_y &= -i\hbar\left(\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right) \\
L_z &= -i\hbar\frac{\partial}{\partial\phi}, \\
\mathbf{L}^2 &= -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right].
\end{aligned} \tag{7.20}$$

The eigenvalue equation for L_z then becomes

$$L_z\psi(r,\theta,\phi) = -i\hbar\frac{\partial}{\partial\phi}\psi(r,\theta,\phi) = m\hbar\psi(r,\theta,\phi) \tag{7.21}$$

We can solve this differential equation to find that

$$\psi(r,\theta,\phi) = \zeta(r,\theta)e^{im\phi}. \tag{7.22}$$

A spatial rotation over 2π must return the wave function to its original value, because $\psi(r,\theta,\phi)$ must have a unique value at each point in space. This leads to $\psi(r,\theta,\phi+2\pi) = \psi(r,\theta,\phi)$ and

$$e^{im(\phi+2\pi)} = e^{im\phi}, \quad \text{or} \quad e^{2\pi im} = 1 \tag{7.23}$$

This means that m is an integer, which in turn means that l must be an integer also.

This page titled 7.1: Orbital Angular Momentum is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

7.2: Spin

For orbital angular momentum we found that $2l+1$ must be an integer, and moreover the spatial properties of the wave function force l to be an integer as well. However, we can also construct states with half-integer l, but this must then be an internal degree of freedom. This is called spin angular momentum, or spin for short. We will show later in the course that the spin observable is interpreted as an intrinsic magnetic moment of a system.

To describe spin, we switch from **L** to **S**, which is no longer related to **r** and **p**. The commutation relations between the components S_i are the same as for L_i,

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k \tag{7.24}$$

so **S** and **L** obey the same algebra. The commutation relations between **S** and **L**, **r**, and **p** vanish:

$$[S_i, L_j] = [S_i, r_j] = [S_i, p_j] = 0. \tag{7.25}$$

Therefore, spin generates a whole new vector space, since it commutes with observables that themselves do not commute (like r_j and p_j), and it is independent of the spatial degrees of freedom.

Since the commutation relations for **S** (its algebra) are the same as for **L**, we can immediately copy the algebraic structure of the eigenstates and eigenvalues:

$$\begin{aligned} S_z \left|s, m_s\right\rangle &= m_s\hbar \left|s, m_s\right\rangle, \quad \text{with } s = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots \\ \mathbf{S}^2 \left|s, m_s\right\rangle &= s(s+1)\hbar^2 \left|s, m_s\right\rangle. \end{aligned} \tag{7.26}$$

When $s = \frac{1}{2}$, the system has two levels (a qubit) with spin eigenstates $\left|\frac{1}{2}, \frac{1}{2}\right\rangle$ and $\left|\frac{1}{2}, -\frac{1}{2}\right\rangle$. We often write $m_s = +\frac{1}{2} = \uparrow$ ("up") and $m_s = -\frac{1}{2} = \downarrow$ ("down"), which finds its origin in the measurement outcomes of electron spin in a Stern-Gerlach apparatus.

Now that we have introduced a whole new vector space related to spin, how do we write the wave function of a particle with spin? Without spin, the wave function is a normal single-valued function $\psi(\mathbf{r}, t) = \langle \mathbf{r} \mid \psi(t) \rangle$ of space and time coordinates. Now we have to add the spin degree of freedom. For each spin $\left(\uparrow \text{ or } \downarrow \text{ when } s = \frac{1}{2}\right)$, we have a wave function $\psi_\uparrow(\mathbf{r}, t)$ for the particle with spin up, and $\psi_\downarrow(\mathbf{r}, t)$ for the particle with spin down. We can write this as a vector:

$$\psi(\mathbf{r}, t) = \begin{pmatrix} \psi_\uparrow(\mathbf{r}, t) \\ \psi_\downarrow(\mathbf{r}, t) \end{pmatrix}. \tag{7.27}$$

The spin degree of freedom generates a vector space, after all. The vector ψ is called a spinor.

Expectation values are evaluated in the usual way, but now we have to sum over the spin degree of freedom, as well as integrate over space. For example, the probability of finding a particle with spin up in a region Ω of space is given by

$$p(\uparrow, \Omega) = \sum_{m_s=\uparrow,\downarrow} \int_\Omega d\mathbf{r} \delta_{m_s,\uparrow} |\psi_{m_s}(\mathbf{r}, t)|^2 = \int_\Omega d\mathbf{r} |\psi_\uparrow(\mathbf{r}, t)|^2, \tag{7,28}$$

and the expectation value of finding a particle with any spin in a region Ω of space is given by

$$p(\Omega) = \sum_{m_s=\uparrow,\downarrow} \int_\Omega d\mathbf{r} |\psi_{m_s}(\mathbf{r}, t)|^2. \tag{7.29}$$

The normalization of the spinor is such that

$$\sum_{m_s=\uparrow,\downarrow} \int_V d\mathbf{r} |\psi_{m_s}(\mathbf{r}, t)|^2 = 1, \tag{7.30}$$

where V is the entire space available to the particle (this may be the entire universe, or the volume of a box with impenetrable walls, etc.).

If spin is represented by $(2s+1)$-dimensional spinors (vectors), then spin transformations (operators) are represented by $(2s+1) \times (2s+1)$ matrices. In the two-dimensional case, we have by construction:

$$S_z|\uparrow\rangle = \frac{\hbar}{2}\begin{pmatrix}1\\0\end{pmatrix} \quad \text{and} \quad S_z|\downarrow\rangle = -\frac{\hbar}{2}\begin{pmatrix}0\\1\end{pmatrix}, \tag{7.31}$$

which means that the matrix representation of S_z is given by

$$S_z = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}. \tag{7.32}$$

Next, the ladder operators act according to

$$S_+|\uparrow\rangle = 0, \quad S_+|\downarrow\rangle = \hbar|\uparrow\rangle, \quad S_-|\uparrow\rangle = \hbar|\downarrow\rangle, \quad S_-|\downarrow\rangle = 0, \tag{7.33}$$

which leads to the matrix representation

$$S_+ = \hbar\begin{pmatrix}0 & 1\\0 & 0\end{pmatrix} \quad \text{and} \quad S_- = \hbar\begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}. \tag{7.34}$$

From $S_\pm = S_x \pm iS_y$ we can then deduce that

$$S_x = \frac{\hbar}{2}\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix} \quad \text{and} \quad S_y = \frac{\hbar}{2}\begin{pmatrix}0 & -i\\i & 0\end{pmatrix}. \tag{7.35}$$

We often define $S_i \equiv \frac{1}{2}\hbar\sigma_i$, where σ_i are the so-called Pauli matrices. Previously, we have called these matrices X, Y, and Z. The commutation relations of the Pauli matrices are

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \quad \text{or} \quad \left[\frac{\sigma_i}{2}, \frac{\sigma_j}{2}\right] = i\epsilon_{ijk}\frac{\sigma_k}{2}. \tag{7.36}$$

Other important properties of the Pauli matrices are

$$\{\sigma_i, \sigma_j\} \equiv \sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}\mathbb{I} \quad \text{(anti-commutator)}. \tag{7.37}$$

They are both Hermitian and unitary, and the square of the Pauli matrices is the identity: $\sigma_i^2 = \mathbb{I}$. Moreover, they obey an "orthogonality" relation

$$\frac{1}{2}\text{Tr}(\sigma_i\sigma_j) = \delta_{ij}. \tag{7.38}$$

The proof of this statement is as follows:

$$\sigma_i\sigma_j = \sigma_i\sigma_j + \sigma_j\sigma_i - \sigma_j\sigma_i = \{\sigma_i, \sigma_j\} - \sigma_j\sigma_i = 2\delta_{ij}^{\mathbb{I}} - \sigma_j\sigma_i. \tag{7.39}$$

Taking the trace then yields

$$\text{Tr}(2\delta_{ij}\mathbb{I} - \sigma_j\sigma_i) = \text{Tr}(\sigma_i\sigma_j) = \text{Tr}(\sigma_j\sigma_i), \tag{7.40}$$

or (using $\text{Tr}(\mathbb{I}) = 2$)

$$2\,\text{Tr}(\sigma_j\sigma_i) = 4\delta_{ij}, \tag{7.41}$$

which proves Eq. (7.38). If we define $\sigma_0 \equiv \mathbb{I}$, we can extend this proof to the four-dimensional case

$$\frac{1}{2}\text{Tr}(\sigma_\mu\sigma_v) = \delta_{\mu v} \tag{7.42}$$

with $\mu, v = 0, 1, 2, 3$.

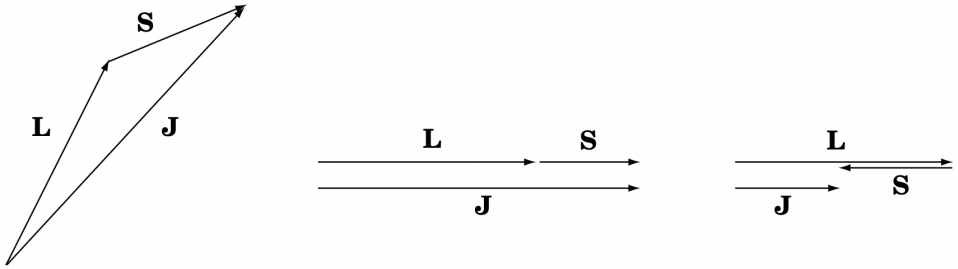


Figure 3: Addition of angular momentum.

We can then write any 2×2 matrix as a sum over the two-dimensional Pauli operators:

$$A=\sum_{\mu} a_{\mu}\sigma_{\mu}, \tag{7.43}$$

since

\(\frac{1}{2} \operatorname{Tr}\left(A \sigma_{v}\right)=\frac{1}{2} \operatorname{Tr}\left(\sum_{\mu} a_{\mu} \sigma_{\mu} \sigma_{v}\right)=\frac{1}{2} \sum_{\mu} a_{\mu} \operatorname{Tr}\left(\sigma_{\mu} \sigma_{v}\right)=\sum_{\mu} a_{\mu} \delta_{\mu v}=a_{v}\tag{7.44}\]

The Pauli matrices and the identity matrix form a basis for the 2×2 matrices, and we can write

$$A=a_0\mathbb{I}+\mathbf{a}\cdot\sigma=\begin{pmatrix} a_0+a_z & a_x-ia_y \\ a_x+ia_y & a_0-a_z \end{pmatrix} \tag{7.45}$$

where we used the notation $\boldsymbol{\sigma}=(\sigma_x,\sigma_y,\sigma_z)$.

This page titled 7.2: Spin is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

7.3: Total Angular Momentum

In general, a particle may have both spin and orbital angular momentum. Since $\mathbf{L}$ and $\mathbf{S}$ have the same dimensions, we can ask what is the total angular momentum $\mathbf{J}$ of the particle. We write this as

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \equiv \mathbf{L} \otimes \mathbb{I} + \mathbb{I} \otimes \mathbf{S}, \tag{7.46}$$

which emphasizes that orbital and spin angular momentum are described in distinct Hilbert spaces.

Since $[L_i, S_j] = 0$, we have

$$\begin{aligned} [J_i, J_j] &= [L_i + S_i, L_j + S_j] = [L_i, L_j] + [S_i, S_j] \\ &= i\hbar\epsilon_{ijk} L_k + i\hbar\epsilon_{ijk} S_k = i\hbar\epsilon_{ijk} \left(L_k + S_k\right) \\ &= i\hbar\epsilon_{ijk} J_k \end{aligned} \tag{7.47}$$

In other words, $\mathbf{J}$ obeys the same algebra as $\mathbf{L}$ and $\mathbf{S}$, and we can immediately carry over the structure of the eigenvalues and eigenvectors from $\mathbf{L}$ and $\mathbf{S}$.

In addition, $\mathbf{L}$ and $\mathbf{S}$ must be added as vectors. However, only one of the components of the total angular momentum can be sharp (i.e., having a definite value). Recall that l and s are magnitudes of the orbital and spin angular momentum, respectively. We can determine the extremal values of $\mathbf{J}$, denoted by $\pm j$, by adding and subtracting the spin from the orbital angular momentum, as shown in Figure 3:

$$|l - s| \leq j \leq l + s. \tag{7.48}$$

For example, when $l = 1$ and $s = \frac{1}{2}$, the possible values of j are $j = \frac{1}{2}$ and $j = \frac{3}{2}$.

The commuting operators for $\mathbf{J}$ are, first of all, $\mathbf{J}^2$ and J_z as we expect from the algebra, but also the operators $\mathbf{L}^2$ and $\mathbf{S}^2$. You may think that S_z and L_z also commute with these operators, but that it not the case:

$$\left[\mathbf{J}^2, L_z\right] = \left[(\mathbf{L} + \mathbf{S})^2, L_z\right] = \left[\mathbf{L}^2 + 2\mathbf{L}\cdot\mathbf{S} + \mathbf{S}^2, L_z\right] = 2\left[\mathbf{L}, L_z\right]\cdot\mathbf{S} \neq 0 \tag{7.49}$$

We can construct a full basis for total angular momentum in terms of $\mathbf{J}^2$ and J_z, as before:

$$\mathbf{J}^2 \left|j, m_j\right\rangle = \hbar^2 j(j+1) \left|j, m_j\right\rangle \text{ and } J_z \left|j, m_j\right\rangle = m_j \hbar \left|j, m_j\right\rangle. \tag{7.50}$$

Alternatively, we can construct spin and orbital angular momentum eigenstates directly as a tensor product of the eigenstates

$$\mathbf{L}^2 |l, m\rangle \left|s, m_s\right\rangle = \hbar^2 l(l+1) |l, m\rangle \left|s, m_s\right\rangle \quad \text{and} \quad L_z |l, m\rangle \left|s, m_s\right\rangle = m\hbar |l, m\rangle \left|s, m_s\right\rangle, \tag{7.51}$$

and

$$\mathbf{S}^2 |l, m\rangle \left|s, m_s\right\rangle = \hbar^2 s(s+1) |l, m\rangle \left|s, m_s\right\rangle \quad \text{and} \quad S_z |l, m\rangle \left|s, m_s\right\rangle = m_s \hbar |l, m\rangle \left|s, m_s\right\rangle. \tag{7.52}$$

Since the L_z and S_z do not commute with $\mathbf{J}^2$, the states $\left|j, m_j\right\rangle$ are not the same as the states $|l, m\rangle \left|s, m_s\right\rangle$.

This page titled 7.3: Total Angular Momentum is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

7.4: Composite Systems with Angular Momentum

Now consider two systems, 1 and 2, with total angular momentum $\mathbf{J}_1$ and J_2, respectively. The total angular momentum is again additive, and given by

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 \equiv \mathbf{J}_1 \otimes \mathbb{I} + \mathbb{I} \otimes \mathbf{J}_2. \tag{7.53}$$

Completely analogous to the addition of spin and orbital angular momentum, we can construct the commuting operators J^2, J_z, $\mathbf{J}_1^2$, and J_2^2, but not J_{1z} and J_{2z}. Again, we construct two natural bases for the total angular momentum of the composite system, namely

$$\{|j,m\rangle\} \quad \text{or} \quad \{|j_1,m_1\rangle \otimes |j_2,m_2\rangle\} \equiv \{|j_1,j_2,m_1,m_2\rangle\} \tag{7.54}$$

We want to know how the two bases relate to each other, because sometimes we wish to talk about the angular momentum of the composite system, and at other times we are interested in the angular momentum of the subsystems. Since the second basis (as well as the first) in Eq. (7.54) forms a complete orthonormal basis, we can write

$$|j,m\rangle = \sum_{m_1,m_2} |j_1,j_2,m_1,m_2\rangle \langle j_1,j_2,m_1,m_2 \mid j,m\rangle \tag{7.55}$$

The amplitudes $\langle j_1,j_2,m_1,m_2 \mid j,m\rangle$ are called Clebsch-Gordan coefficients, and we will now present a general procedure for calculating them.

Let's first consider a simple example of two spin$-\frac{1}{2}$ systems, such as two electrons. The spin basis for each electron is given by $\left|\frac{1}{2},\frac{1}{2}\right\rangle = |\uparrow\rangle$ and $\left|\frac{1}{2},-\frac{1}{2}\right\rangle = |\downarrow\rangle$. The spin basis for the two electrons is therefore

$$|j_1,j_2,m_1,m_2\rangle \in \{|\uparrow,\uparrow\rangle, |\uparrow,\downarrow\rangle, |\downarrow,\uparrow\rangle, |\downarrow,\downarrow\rangle\}. \tag{7.56}$$

The total spin is given by $j = \frac{1}{2} + \frac{1}{2} = 1$ and $j = \frac{1}{2} - \frac{1}{2} = 0$, so the four basis states for total angular momentum are

$$|j,m\rangle \in \{|1,1\rangle, |1,0\rangle, |1,-1\rangle, |0,0\rangle\} \tag{7.57}$$

The latter state is the eigenstate for $j=0$. The maximum total angular momentum state $|1,1\rangle$ can occur only when the two electron spins a parallel, and we therefore have

$$|1,1\rangle = |\uparrow,\uparrow\rangle = \left|\frac{1}{2},\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2},\frac{1}{2}\right\rangle \tag{7.58}$$

To find the expansion of the other total angular momentum eigenstates in terms of spin eigenstates we employ the following trick: use that $J_\pm = J_{1\pm} + J_{2\pm}$. We can then apply $J_\pm$ to the state $|1,1\rangle$, and $J_{1\pm} + J_{2\pm}$ to the state $\left|\frac{1}{2},\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2},\frac{1}{2}\right\rangle$. This yields

$$J_- |1,1\rangle = \hbar\sqrt{j(j+1) - m(m-1)}\, |1,0\rangle = \hbar\sqrt{2}|1,0\rangle \tag{7.59}$$

Similarly, we calculate

$$J_{1-}\left|\frac{1}{2},\frac{1}{2}\right\rangle = \hbar\sqrt{\frac{1}{2}\left(\frac{3}{2}\right) - \frac{1}{2}\left(-\frac{1}{2}\right)}\left|\frac{1}{2},-\frac{1}{2}\right\rangle = \hbar\left|\frac{1}{2},-\frac{1}{2}\right\rangle, \tag{7.60}$$

and a similar result for J_{2-}. Therefore, we find that

$$\hbar\sqrt{2}|1,0\rangle = \hbar|\uparrow,\downarrow\rangle + \hbar|\downarrow,\uparrow\rangle \quad \Longrightarrow \quad |1,0\rangle = \frac{|\uparrow,\downarrow\rangle + |\downarrow,\uparrow\rangle}{\sqrt{2}}. \tag{7.61}$$

Applying J_- again yields

$$|1,-1\rangle = |\downarrow,\downarrow\rangle \tag{7.62}$$

This agrees with the construction of adding parallel spins. The three total angular momentum states

$$
\begin{aligned}
|1,1\rangle &= |\uparrow,\uparrow\rangle, \\
|1,0\rangle &= \frac{1}{\sqrt{2}}(|\uparrow,\downarrow\rangle + |\downarrow,\uparrow\rangle), \\
|1,-1\rangle &= |\downarrow,\downarrow\rangle
\end{aligned}
\tag{7.63}
$$

form a so-called triplet of states with $j=1$. We now have to find the final state corresponding to $j=0, m=0$. The easiest way to find it at this point is to require orthonormality of the four basis states, and this gives us the singlet state

$$
|0,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle). \tag{7.64}
$$

The singlet state has zero total angular momentum, and it is therefore invariant under rotations.

In general, this procedure of finding the Clebsch-Gordan coefficients results in multiplets of constant j. In the case of two spins, we have a tensor product of two two-dimensional spaces, which are decomposed in two subspaces of dimension 3 (the triplet) and 1 (the singlet), respectively. We write this symbolically as

$$
2 \otimes 2 = 3 \oplus 1. \tag{7.65}
$$

If we had combined a spin 1 particle with a spin $\frac{1}{2}$ particle, the largest multiplet would have been due to $j = 1 + \frac{1}{2} = \frac{3}{2}$, which is a 4-dimensional subspace, and the smallest subspace is due to $j = 1 - \frac{1}{2} = \frac{1}{2}$, which is a two-dimensional subspace:

$$
3 \otimes 2 = 4 \oplus 2. \tag{7.66}
$$

In general, the total angular momentum of two systems with angular momentum k and l is decomposed into multiplets according to the following rule $(k \geq l)$:

$$
(2k+1) \otimes (2l+1) = [2(k+l)+1] \oplus [2(k+l)-1] \oplus \ldots \oplus [2(k-l)+1], \tag{7.67}
$$

or in terms of the dimensions of the subspaces $(n \geq m)$:

$$
n \otimes m = (n+m-1) \oplus (n+m-3) \oplus \ldots \oplus (n-m+1). \tag{7.68}
$$

? Exercises

1. Angular momentum algebra.
   a. Prove the algebra given in Eq. (7.2). Also show that $\left[L^2, L_i\right] = 0$, and verify the commutation relations in Eq. (7.7).
   b. Show that $L^2|l,m\rangle = l(l+1)\hbar^2|l,m\rangle$. Use the fact that $\left[L_-, L^2\right] = 0$.
2. Pauli matrices.
   a. Check that the matrix representation of the spin$-\frac{1}{2}$ operators obey the commutation relations.
   b. Calculate the matrix representation of the Pauli matrices for $s=1$.
   c. Prove that $\exp[-i\theta\cdot\sigma]$ is a 2×2 unitary matrix.
3. Isospin $\mathbf{I}$ describes certain particle families called multiplets, and the components of the isospin obey the commutation relations $[I_i, I_j] = i\epsilon_{ijk}I_k$.
   a. What is the relation between spin and isospin?
   b. Organize the nucleons (proton and neutron), the pions $\left(\pi^+, \pi^0, \text{ and } \pi^-\right)$, and the delta baryons $\left(\Delta^{++}, \Delta^+, \Delta^0, \text{ and } \Delta^-\right)$ into multiplets. You will have to determine their isospin quantum number.
   c. Give all possible decay channels of the delta baryons into pions and nucleons (use charge and baryon number conservation).
   d. Calculate the relative decay ratios of Δ^+ and Δ^0 into the different channels.
4. A simple atom has orbital and spin angular momentum, and the Hamiltonian for the atom contains a spin-orbit coupling term $H_{\text{so}} = g\hbar\mathbf{L}\cdot\mathbf{S}$, where $g\hbar$ is the coupling strength.
   a. Are orbital and spin angular momentum good quantum numbers for this system? What about total angular momentum?
   b. Use first-order perturbation theory to calculate the energy shift due to the spin-orbit coupling term.
   c. Calculate the transition matrix elements of H_{so} in the basis $\{|l,m;s,m_s\rangle\}$.
5. Multiplets.

a. A spin $\frac{3}{2}$ particle and a spin 2 particle form a composite system. How many multiplets are there, and what is the dimension of the largest multiplet?
b. How many multiplets do two systems with equal angular momentum have?

This page titled 7.4: Composite Systems with Angular Momentum is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

8: Identical Particles

We have so far looked at the quantum mechanical description of a few particles with spin in the previous section, and particles that exhibit entanglement in section 5. In all these cases, we assumed that the individual particles could be distinguished from each other. For example, the two-electron state $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$ assumes that we have two electrons, one held "over here", and the other "over there", and we can talk meaningfully about their respective spins. The tensor product structure of our Hilbert space is a manifestation of our ability to label our particles unambiguously.

However, what happens when we place the two electrons inside a sealed box? The wave functions of the electrons will quickly start to overlap. Since the electrons are identical particles, which according to basic quantum mechanics do not have well-defined paths, we cannot keep track of which electron is which inside the box. Not even in principle.

8.1: Symmetric and Anti-symmetric States

8.2: Creation and Annihilation Operators

8.3: Observables Based on Creation and Annihilation Operators

8.4: Bose-Einstein and Fermi-Dirac Statistics

This page titled 8: Identical Particles is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

8.1: Symmetric and Anti-symmetric States

The indistinguishability of identical particles means that we have to adjust our quantum mechanical description of these objects. There are two ways of doing this, namely via a modification of the allowed states and via a restructuring of the observables[5]. In this section we consider the restricted state space, and in the next we will be considering the new observables.

First of all, since the total number of particles is an observable quantity (for example by measuring the total charge in the box), we can give the particles an artificial labelling. The wave functions of the two particles are then given by $|\psi(\mathbf{r}_1)\rangle_1$ for particle 1 at position $\boldsymbol{r}_1$, and $|\phi(\mathbf{r}_2)\rangle_2$ for particle 2 at position $\mathbf{r}_2$. Since we can swap the positions of the particle without observable consequences, we find that there are two states that denote the same physical situation:

$$|\psi(\mathbf{r}_1),\phi(\mathbf{r}_2)\rangle_{12} \quad \text{and} \quad |\psi(\mathbf{r}_2),\phi(\mathbf{r}_1)\rangle_{12}. \tag{8.1}$$

However, we wish that each physically distinct situation has exactly one quantum state. Since there is no preference for either state, we can denote the physical situation of identical particles at position $\mathbf{r}_1$ and $\mathbf{r}_2$ by the quantum state that is an equal weight over these two possibilities:

$$|\Psi(\mathbf{r}_1,\mathbf{r}_2)\rangle_{12} = \frac{|\psi(\mathbf{r}_1),\phi(\mathbf{r}_2)\rangle_{12} + e^{i\varphi}|\psi(\mathbf{r}_2),\phi(\mathbf{r}_1)\rangle_{12}}{\sqrt{2}}. \tag{8.2}$$

You can verify that swapping the positions $\mathbf{r}_1$ and $\mathbf{r}_2$ of the indistinguishable particles incurs only a global (unobservable) phase. The question is now how we should choose ϕ.

Suppose that the two identical particles in the box are electrons. We know From Pauli's exclusion principle that the two electrons cannot be in the same state. Therefore, when $\phi = \psi$, the state in Eq. (8.2) should naturally disappear:

$$|\Psi(\mathbf{r}_1,\mathbf{r}_2)\rangle_{12} = \frac{|\psi(\mathbf{r}_1),\psi(\mathbf{r}_2)\rangle_{12} + e^{i\varphi}|\psi(\mathbf{r}_2),\psi(\mathbf{r}_1)\rangle_{12}}{\sqrt{2}} = 0, \tag{8.3}$$

which means that for particles obeying Pauli's exclusion principle we must choose $e^{i\varphi} = -1$. The quantum state of the two particles is anti-symmetric.

What about particles that do not obey Pauli's exclusion principle? These must be restricted to states that are orthogonal to the anti-symmetric states. In other words, they must be in states that are symmetric under the exchange of two particles. For the two identical particles in a box, we therefore choose the value $e^{i\varphi} = +1$, which makes the state orthogonal to the anti-symmetric state. The two possibilities for combining two identical particles are therefore

$$\begin{aligned} |\Psi_{\mathrm{S}}(\mathbf{r}_1,\mathbf{r}_2)\rangle &= \frac{|\psi(\mathbf{r}_1),\phi(\mathbf{r}_2)\rangle + |\psi(\mathbf{r}_2),\phi(\mathbf{r}_1)\rangle}{\sqrt{2}} \\ |\Psi_{\mathrm{A}}(\mathbf{r}_1,\mathbf{r}_2)\rangle &= \frac{|\psi(\mathbf{r}_1),\phi(\mathbf{r}_2)\rangle - |\psi(\mathbf{r}_2),\phi(\mathbf{r}_1)\rangle}{\sqrt{2}} \end{aligned} \tag{8.4}$$

These states include both the internal degrees of freedom, such as spin, and the external degrees of freedom. So two electrons can still be in the state $|\uparrow\uparrow\rangle$, as long as their spatial wave function is anti-symmetric. The particles that are in a symmetric overall quantum state are bosons, while the particles in an overall anti-symmetric state are fermions.

We can extend this to N particles in a fairly straightforward manner. For bosons, we sum over all possible permutations of $\mathbf{r}_1$ to $\mathbf{r}_N$:

$$|\Psi_{\mathrm{S}}(\mathbf{r}_1,\ldots,\mathbf{r}_N)\rangle = \frac{1}{\sqrt{N!}} \sum_{\mathrm{perm}(\mathbf{r}_1,\ldots,\mathbf{r}_N)} |\psi_1(\mathbf{r}_1),\ldots,\psi_N(\mathbf{r}_N)\rangle. \tag{8.5}$$

For fermions, the odd permutations pick up a relative minus sign:

$$|\Psi_{\mathrm{A}}(\mathbf{r}_1,\ldots,\mathbf{r}_N)\rangle = \frac{1}{\sqrt{N!}} \sum_{\text{even}} |\psi_1(\mathbf{r}_1),\ldots,\psi_N(\mathbf{r}_N)\rangle - \frac{1}{\sqrt{N!}} \sum_{\text{odd}} |\psi_1(\mathbf{r}_1),\ldots,\psi_N(\mathbf{r}_N)\rangle \tag{8.6}$$

This can be written compactly as the so-called **Slater determinant**

$$\Psi_{\mathrm{A}}\left(\mathbf{r}_1,\ldots,\mathbf{r}_N\right)=\frac{1}{\sqrt{N!}}\begin{vmatrix}\psi_1\left(\mathbf{r}_1\right) & \psi_1\left(\mathbf{r}_2\right) & \ldots & \psi_1\left(\mathbf{r}_N\right)\\ \psi_2\left(\mathbf{r}_1\right) & \psi_2\left(\mathbf{r}_2\right) & \ldots & \psi_2\left(\mathbf{r}_N\right)\\ \vdots & \vdots & \ddots & \vdots\\ \psi_N\left(\mathbf{r}_1\right) & \psi_N\left(\mathbf{r}_2\right) & \ldots & \psi_N\left(\mathbf{r}_N\right)\end{vmatrix}, \tag{8.7}$$

where we removed the kets for notational convenience. The N particles in the state $\left|\Psi_{\mathrm{A}}\left(\mathbf{r}_1,\ldots,\mathbf{r}_N\right)\right\rangle$ automatically obey the Pauli exclusion principle.

[5]This is sometimes called second quantisation. This is a misnomer, since quantisation occurs only once, when observables are promoted to operators.

This page titled 8.1: Symmetric and Anti-symmetric States is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

8.2: Creation and Annihilation Operators

The second, and particularly powerful way to implement the description of identical particles is via creation and annihilation operators. To see how this description arises, consider some single-particle Hermitian operator A with eigenvalues a_j. On physical grounds, and regardless of distinguishability, we require that n_j particles in the eigenstate $|a_j\rangle$ of A must have a total physical value $n_j \times a_j$ for the observable A. We can repeat this for all eigenvalues a_j, and obtain a potentially infinite set of basis vectors

$$|n_1, n_2, n_3, \ldots\rangle ,$$

for all integer values of n_j, including zero. You should convince yourself that this exhausts all the possible ways any number of particles can be distributed over the eigenvalues a_j. The spectrum of A can be bounded or unbounded, and discrete or continuous. It may even be degenerate. For simplicity we consider here an unbounded, non-degenerate discrete spectrum.

A special state is given by

$$|\varnothing\rangle = |0, 0, 0, \ldots\rangle \tag{8.2.1}$$

which indicates the state of no particles, or the vacuum. The numbers n_j are called the occupation number, and any physical state can be written as a superposition of these states:

$$|\Psi\rangle = \sum_{n_1,n_2,n_3,\ldots=0}^{\infty} c_{n_1,n_2,n_3,\ldots} |n_1, n_2, n_3, \ldots\rangle . \tag{8.2.2}$$

The basis states $|n_1, n_2, n_3, \ldots\rangle$ span a linear vector space called a **Fock space** $\mathscr{F}$. It is the direct sum of the Hilbert spaces for zero particles $\mathscr{H}_0$, one particle $\mathscr{H}_1$, two particles, etc.:

$$\mathscr{F} = \mathscr{H}_0 \oplus \mathscr{H}_1 \oplus \mathscr{H}_2 \oplus \mathscr{H}_3 \oplus \cdots \tag{8.2.3}$$

Since $|\Psi\rangle$ is now a superposition over different particle numbers, we require operators that change the particle number. These are the **creation** and **annihilation** operators, $\hat{a}^\dagger$ and $\hat{a}$ respectively. Up to a proportionality constant that we will determine later, the action of these operators is defined by

$$\hat{a}_j^\dagger |n_1, n_2, \ldots, n_j, \ldots\rangle \propto |n_1, n_2, \ldots, n_j + 1, \ldots\rangle$$
$$\hat{a}_j |n_1, n_2, \ldots, n_j, \ldots\rangle \propto |n_1, n_2, \ldots, n_j - 1, \ldots\rangle$$

So the operator $\hat{a}_j^\dagger$ creates a particle in a state with eigenvalue a_j, and the operator $\hat{a}_j$ removes a particle in a state with eigenvalue a_j. These operators are each others' Hermitian adjoint, since removing a particle is the time reversal of adding a particle. Clearly, when an annihilation operator attempts to remove particles that are not there, the result must be zero:

$$\hat{a}_j |n_1, n_2, \ldots, n_j = 0, \ldots\rangle = 0 \tag{8.2.4}$$

The vacuum is then defined as the state that gives zero when acted on by any annihilation operator: $\hat{a}_j|\varnothing\rangle = 0$ for any j. Notice how we have so far sidestepped the problem of particle swapping; we exclusively used aspects of the total particle number.

What are the basic properties of these creation and annihilation operators? In particular, we are interested in their commutation relations. We will now derive these properties from what we have determined so far. First, note that we can create two particles with eigenvalues a_i and a_j in the system in any order, and the only difference this can make is in the normalization of the state:

$$\hat{a}_i^\dagger \hat{a}_j^\dagger |\Psi\rangle = \lambda \hat{a}_j^\dagger \hat{a}_i^\dagger |\Psi\rangle, \tag{8.2.5}$$

where λ is some complex number. Since state $|\Psi\rangle$ is certainly not zero, we require that

$$\hat{a}_k^\dagger \hat{a}_l^\dagger - \lambda \hat{a}_l^\dagger \hat{a}_k^\dagger = 0. \tag{8.2.6}$$

Since k and l are just dummy variables, we equally have

$$\hat{a}_l^\dagger \hat{a}_k^\dagger - \lambda \hat{a}_k^\dagger \hat{a}_l^\dagger = 0. \tag{8.2.7}$$

We now substitute Eq. (8.15) into Eq. (8.14) to eliminate $\hat{a}_l^\dagger \hat{a}_k^\dagger$. This leads to

$$\left(1 - \lambda^2\right) \hat{a}_k^\dagger \hat{a}_l^\dagger = 0, \tag{8.2.8}$$

and therefore

$$\lambda = \pm 1. \tag{8.2.9}$$

The relation between different creation operators can thus take two forms. They can obey a commutation relation when $\lambda = +1$:

$$\hat{a}_l^\dagger \hat{a}_k^\dagger - \hat{a}_k^\dagger \hat{a}_l^\dagger = \left[\hat{a}_l^\dagger, \hat{a}_k^\dagger\right] = 0, \tag{8.2.10}$$

or they can obey an anti-commutation relation when $\lambda = -1$:

$$\hat{a}_l^\dagger \hat{a}_k^\dagger + \hat{a}_k^\dagger \hat{a}_l^\dagger = \left\{\hat{a}_l^\dagger, \hat{a}_k^\dagger\right\} = 0 \tag{8.2.11}$$

While creating the particles in different temporal order is not the same as swapping two particles, it should not come as a surprise that there are two possible situations (the commutation relation and the anti-commutation relation). We encountered two possibilities in our previous approach as well, where we found that many-particle states are either symmetric or anti-symmetric. In fact, creation operators that obey the commutation relation produce symmetric states, while creation operators that obey the anti-commutation relation produce anti-symmetric states. We also see that the creation operators described by the anti-commutation relations naturally obey Pauli's exclusion principle. Suppose that we wish to create two identical particles in the same eigenstate $|a_j\rangle$. The anti-commutation relations say that $\left\{\hat{a}_j^\dagger, \hat{a}_j^\dagger\right\} = 0$, so

$$\hat{a}_j^{\dagger 2} = 0. \tag{8.2.12}$$

Any higher powers of $\hat{a}_j^\dagger$ will also be zero, and we can create at most one particle in the state $|a_j\rangle$.

Taking the adjoint of the commutation relations for the creation operators gives us the corresponding relations for the annihilation operators

$$\hat{a}_l \hat{a}_k - \hat{a}_k \hat{a}_l = [\hat{a}_l, \hat{a}_k] = 0, \tag{8.2.13}$$

or

$$\hat{a}_l \hat{a}_k + \hat{a}_k \hat{a}_l = \{\hat{a}_l, \hat{a}_k\} = 0. \tag{8.2.14}$$

The remaining question is now what the (anti-) commutation relations are for products of creation and annihilation operators.

We proceed along similar lines as before. Consider the operators $\hat{a}_j$ and $\hat{a}_k^\dagger$ with $j \neq k$, and apply them in different orders to a state $|\Psi\rangle$.

$$\hat{a}_i \hat{a}_j^\dagger |\Psi\rangle = \mu \hat{a}_j^\dagger \hat{a}_i |\Psi\rangle. \tag{8.2.15}$$

The same argumentation as before leads to $\mu = \pm 1$. For different j and k we therefore find

$$\left[\hat{a}_j, \hat{a}_k^\dagger\right] = 0 \quad \text{or} \quad \left\{\hat{a}_j, \hat{a}_k^\dagger\right\} = 0 \tag{8.2.16}$$

Now let's consider the case $j = k$. For the special case where $|\Psi\rangle = |\varnothing\rangle$, we find

$$\left(\hat{a}_j \hat{a}_k^\dagger - \mu \hat{a}_j^\dagger \hat{a}_k\right) |\varnothing\rangle = \hat{a}_j \hat{a}_k^\dagger |\varnothing\rangle = \delta_{jk} |\varnothing\rangle, \tag{8.2.17}$$

based on the property that $\hat{a}_j |\varnothing\rangle = 0$. When $l = k$,

$$\left(\hat{a}_k \hat{a}_k^\dagger - \mu \hat{a}_k^\dagger \hat{a}_k\right) |\varnothing\rangle = |\varnothing\rangle, \tag{8.2.18}$$

we find for the two possible values of μ

$$\hat{a}_k \hat{a}_k^\dagger - \hat{a}_k^\dagger \hat{a}_k = 1 \quad \text{or} \quad \hat{a}_k \hat{a}_k^\dagger + \hat{a}_k^\dagger \hat{a}_k = 1 \tag{8.2.19}$$

which is equivalent to

$$\left[\hat{a}_k, \hat{a}_k^\dagger\right] = 1 \quad \text{or} \quad \left\{\hat{a}_k, \hat{a}_k^\dagger\right\} = 1. \tag{8.2.20}$$

To summarise, we have two sets of algebras for the creation and annihilation operators. The algebra in terms of the commutation relations is given by

$$[\hat{a}_k, \hat{a}_l] = \left[\hat{a}_k^\dagger, \hat{a}_l^\dagger\right] = 0 \quad \text{and} \quad \left[\hat{a}_k, \hat{a}_l^\dagger\right] = \delta_{kl}. \tag{8.2.21}$$

This algebra describes particles that obey Bose-Einstein statistics, or bosons. The algebra in terms of anti-commutation relations is given by

$$\{\hat{a}_k, \hat{a}_l\} = \left\{\hat{a}_k^\dagger, \hat{a}_l^\dagger\right\} = 0 \quad \text{and} \quad \left\{\hat{a}_k, \hat{a}_l^\dagger\right\} = \delta_{kl}. \tag{8.2.22}$$

This algebra describes particles that obey Fermi-Dirac statistics, or fermions.

Finally, we have to determine the constant of proportionality for the creation and annihilation operators. We have already required that $\hat{a}_j\hat{a}_k^\dagger|\varnothing\rangle = \delta_{jk}|\varnothing\rangle$. To determine the rest, we consider a new observable that gives us the total number of particles in the system. We denote this observable by $\hat{n}$, and we see that it must be additive over all particle numbers for the different eigenvalues of A:

$$\hat{n} = \sum_j \hat{n}_j, \tag{8.2.23}$$

where $\hat{n}_j$ is the number of particles in the eigenstate $|a_j\rangle$. The total number of particles does not change if we consider a different observable (although the distribution typically will), so this relation is also true when we count the particles in the states $|b_j\rangle$. Pretty much the only way we can achieve this is to choose

$$\hat{n} = \sum_j \hat{n}_j = \sum_j \hat{a}_j^\dagger\hat{a}_j = \sum_j \hat{b}_j^\dagger\hat{b}_j. \tag{8.2.24}$$

For the case of n_j particles in state $|a_j\rangle$ this gives

$$\hat{a}_j^\dagger\hat{a}_j\,|n_j\rangle = n_j\,|n_j\rangle\,, \tag{8.2.25}$$

where we ignored the particles in other states $|a_k\rangle$ with $k \neq j$ for brevity. For the Bose-Einstein case this leads to the relations

$$\hat{a}_j\,|n_j\rangle = \sqrt{n_j}\,|n_j - 1\rangle \quad \text{and} \quad \hat{a}_j^\dagger\,|n_j\rangle = \sqrt{n_j + 1}\,|n_j + 1\rangle\,. \tag{8.2.26}$$

For Fermi-Dirac statistics, the action of the creation and annihilation operators on number states becomes

$$\begin{aligned} \hat{a}_j|0\rangle_j &= 0 \qquad \text{and} \quad \hat{a}_j^\dagger|0\rangle_j = e^{-i\alpha}|1\rangle_j, \\ \hat{a}_j|1\rangle_j &= e^{i\alpha}|0\rangle_j \quad \text{and} \quad \hat{a}_j^\dagger|1\rangle_j = 0. \end{aligned} \tag{8.2.27}$$

The phase factor $e^{i\alpha}$ can be chosen ± 1.

This page titled 8.2: Creation and Annihilation Operators is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

8.3: Observables Based on Creation and Annihilation Operators

So far, we have considered only the basis states of many particles for a single observable A. What about other observables, in particular those that do not commute with A? We can make a similar construction. Suppose an observable B has eigenvalues b_j. We can construct creation and annihilation operators $\hat{b}_j^\dagger$ and $\hat{b}_j$ that act according to

$$\begin{aligned}\hat{b}_j^\dagger |m_1, m_2, \ldots, m_j, \ldots\rangle &= \sqrt{m_j+1}\,|m_1, m_2, \ldots, m_j+1, \ldots\rangle\,,\\ \hat{b}_j |m_1, m_2, \ldots, m_j, \ldots\rangle &= \sqrt{m_j}\,|m_1, m_2, \ldots, m_j-1, \ldots\rangle\,.\end{aligned} \tag{8.36}$$

where m_j is the number of particles with value b_j. Typically, the basis states of two observables are related via a single unitary transformation $|b_j\rangle = U|a_j\rangle$ for all j. How does this relate the creation and annihilation operators?

To answer this, let's look at the single particle states. We can write the single-particle eigenstates $|a_j\rangle$ and $|b_j\rangle$ as

$$|a_j\rangle = \hat{a}_j^\dagger|\varnothing\rangle \quad \text{and} \quad |b_j\rangle = \hat{b}_j^\dagger|\varnothing\rangle. \tag{8.37}$$

We assume that U does not change the vacuum[6], so $U|\varnothing\rangle = |\varnothing\rangle$. This means that we can relate the two eigenstates via

$$|b_j\rangle = U|a_j\rangle = U\hat{a}_j^\dagger|\varnothing\rangle = U\hat{a}_j^\dagger\left(U^\dagger U\right)|\varnothing\rangle = U\hat{a}_j^\dagger U^\dagger|\varnothing\rangle = \hat{b}_j^\dagger|\varnothing\rangle, \tag{8.38}$$

where we have strategically inserted the identity $\mathbb{I} = U^\dagger U$. This leads to the operator transformation

$$\hat{b}_j^\dagger = U\hat{a}_j^\dagger U^\dagger. \tag{8.39}$$

The Hermitian adjoint is easily calculated as $\hat{b}_j = U\hat{a}U^\dagger$. It is left as an exercise for you to prove that

$$\hat{b}_j^\dagger = \sum_k u_{jk}\hat{a}_k^\dagger \quad \text{and} \quad \hat{b}_j = \sum_k u_{kj}^*\hat{a}_k, \tag{8.40}$$

where $u_{jk} = \langle a_k \mid b_j\rangle$.

How do we construct operators using the creation and annihilation operators? Suppose that a one-particle observable H has eigenvalues E_j and eigenstates $|j\rangle$. This can be, for example the Hamiltonian of the system, which ensures that the physical values of the particles (the eigenvalues) are additive. The operator for many identical particles then becomes

$$H = \sum_j E_j\hat{n}_j = \sum_j E_j\hat{a}_j^\dagger\hat{a}_j, \tag{8.41}$$

which transforms according to Eq. (8.40). More generally, the operator may not be written in the eigenbasis $|n_1, n_2, \ldots\rangle$, in which case it has the form

$$H = \sum_{ij} H_{ij}\hat{b}_i^\dagger\hat{b}_j, \tag{8.42}$$

where H_{ij} are matrix elements. The creation and annihilation operators $\hat{a}_j^\dagger$ and $\hat{a}_j$ diagonalise H, and are sometimes called normal modes. The reason for this is that the creation and annihilation operators for bosons obey the same mathematical rules as the raising and lowering operators for the harmonic oscillator. The index j then denotes different oscillators. A system of coupled oscillators can be decomposed into normal modes, which are themselves isolated harmonic oscillators.

[6]This is a natural assumption when we are confined to the single particle Hilbert space, but there are general unitary transformations for which this does not hold, such as the transformation to an accelerated frame.

This page titled 8.3: Observables Based on Creation and Annihilation Operators is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

8.4: Bose-Einstein and Fermi-Dirac Statistics

Finally, in this section we will derive the Bose-Einstein and Fermi-Dirac statistics. In particular, we are interested in the thermal equilibrium for a large number of (non-interacting) identical particles with some energy spectrum E_j, which my be continuous.

Since the number of particles is not fixed, we are dealing with the Grand Canonical Ensemble. Its partition function Ξ is given by

$$\Xi = \mathrm{Tr}\left[e^{\mu\beta\hat{n}-\beta H}\right], \tag{8.43}$$

where H is the many-body Hamiltonian, $\beta = 1/k_B T$ and μ is the chemical potential. The average number of particles with single particle energy $\boldsymbol{E}_j$ is then given by

$$\langle n_j \rangle = -\frac{1}{\beta}\frac{\partial \ln \Xi}{\partial E_j}. \tag{8.44}$$

For the simple case where $H = \sum_j E_j \hat{n}_j$ and the creation and annihilation operators obey the commutator algebra, the exponent can be written as

$$\exp\left[\beta \sum_j (\mu - E_j)\, \hat{a}_j^\dagger \hat{a}_j\right] = \bigotimes_j \sum_{n_j=0}^{\infty} e^{\beta(\mu-E_j)n_j} \left|n_j\right\rangle \left\langle n_j\right|, \tag{8.45}$$

and the trace becomes

$$\Xi = \prod_j \frac{1}{1 - e^{\beta(\mu - E_j)}}. \tag{8.46}$$

The average photon number for energy $\boldsymbol{E}_j$ is

$$\langle n_j \rangle = -\frac{1}{\beta}\frac{\partial \ln \Xi}{\partial E_j} = -\frac{1}{\beta \Xi}\frac{\partial \Xi}{\partial E_j} = \frac{1}{e^{-\beta(\mu - E_j)} - 1}. \tag{8.47}$$

This is the Bose-Einstein distribution for particles with energy E_j. It is shown for increasing E_j in Fig. 4 on the left.

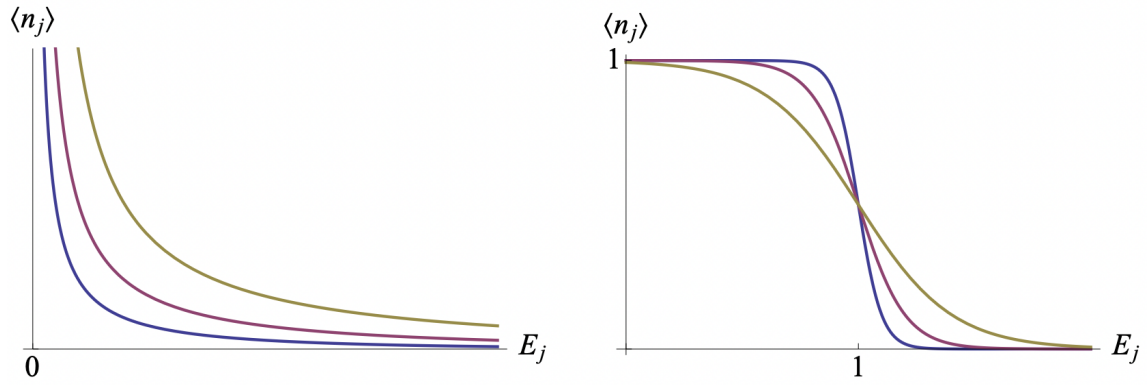


Figure 4: Left: Bose-Einstein distribution for different temperatures ($\mu = 0$). The lower the temperature, the more particles occupy the low energy states. Right: Fermi-Dirac distribution for different temperatures and $\mu = 1$. The fermions will not occupy energy states with numbers higher than 1, and therefore higher energies are necessarily populated. The energy values E_j form a continuum on the horizontal axis.

Alternatively, if the creation and annihilation operators obey the anti-commutation relations, the sum over n_j in Eq. (8.45) runs not from 0 to ∞, but over 0 and 1. The partition function of the grand canonical ensemble then becomes

$$\Xi = \prod_j \left[1 + e^{\beta(\mu - E_j)}\right], \tag{8.48}$$

and the average number of particles with energy E_j becomes

$$\langle n_j \rangle = -\frac{1}{\beta \Xi}\frac{\partial \Xi}{\partial \hbar\omega_j} = \frac{1}{e^{-\beta(\mu - E_j)} + 1}. \tag{8.49}$$

This is the Fermi-Dirac statistics for these particles, and it is shown in Fig. 4 on the right. The chemical potential is the highest occupied energy at zero temperature, and in solid state physics this is called the Fermi level. Note the sign difference in the denominator with respect to the Bose-Einstein statistics.

Exercises

1. Calculate the Slater determinant for three electrons and show that no two electrons can be in the same state.
2. Particle statistics.
   a. What is the probability of finding n bosons with energy E_j in a thermal state?
   b. What is the probability of finding n fermions with energy E_j in a thermal state?
3. Consider a system of (non-interacting) identical bosons with a discrete energy spectrum and a ground state energy $\boldsymbol{E}_0$. Furthermore, the chemical potential starts out lower than the ground state energy $\mu < E_0$.
   a. Calculate $\langle n_0 \rangle$ and increase the chemical potential to $\mu \to E_0$ (e.g., by lowering the temperature). What happens when μ passes $\boldsymbol{E}_0$?
   b. What is the behaviour of $\langle n_{\rm thermal} \rangle \equiv \sum_{j=1}^{\infty} \langle n_j \rangle$ as $\mu \to E_0$? Sketch both $\langle n_0 \rangle$ and $\langle n_{\rm thermal} \rangle$ as a function of μ. What is the fraction of particles in the ground state at $\mu = E_0$?
   c. What physical process does this describe?
4. The process $U = \exp\left(r\hat{a}_1^\dagger \hat{a}_2^\dagger - r^* \hat{a}_1 \hat{a}_2\right)$ with $r \in \mathbb{C}$ creates particles in two systems, 1 and 2, when applied to the vacuum state $|\Psi\rangle = U|\varnothing\rangle$.
   a. Show that the bosonic operators $\hat{a}_1^\dagger \hat{a}_2^\dagger$ and $\hat{a}_1 \hat{a}_2$ obey the algebra
   $$[K_-, K_+] = 2K_0 \quad \text{and} \quad [K_0, K_\pm] = \pm K_\pm,$$
   with $K_+ = K_-^\dagger$.
   b. For operators obeying the algebra in (a) we can write
   $$e^{rK_+ - r^* K_-} = \exp\left[\frac{r}{|r|}\tanh|r| K_+\right] \exp[-2\ln(\cosh|r|) K_0] \times \exp\left[-\frac{r^*}{|r|}\tanh|r| K_-\right]. \tag{8.50}$$
   Calculate the state $|\Psi\rangle$ of the two systems.
   c. The amount of entanglement between two systems can be measured by the entropy $S(r)$ of the reduced density matrix $\rho_1 = \mathrm{Tr}_1[\rho]$ for one of the systems. Calculate $S(r) = -\mathrm{Tr}[\rho_1 \ln \rho_1]$.
   d. What is the probability of finding n particles in system 1?

This page titled 8.4: Bose-Einstein and Fermi-Dirac Statistics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

9: Many-Body Problems in Quantum Mechanics

In this final section we study a variety of problems in many-body quantum mechanics. First, we introduce the Hartree-Fock method for taking into account the effect of electron-electron interactions in atoms. Next, we describe spin waves in magnetic materials using the Heisenberg model. Third, we describe the behaviour of an atom interacting with photons in a cavity, and introduce the Jaynes-Cummings Hamiltonian. And finally, we take a brief look at the basic ideas behind quantum field theory.

9.1: Interacting Electrons in Atomic Shells

9.2 Spin Waves in Solids

9.3: An Atom in a Cavity

9.4: Outlook: Quantum Field Theory

This page titled 9: Many-Body Problems in Quantum Mechanics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

9.1: Interacting Electrons in Atomic Shells

Previously, you have encountered the Schrödinger equation for a particle in a central potential, which can be interpreted as an electron bound to a proton in a hydrogen atom. We found that the energy levels are quantised with quantum number n. In addition, the spin and orbital angular momentum is quantised with quantum numbers m_s, l, and m, respectively (we assume that $s = \frac{1}{2}$ since we are considering electrons). We can denote the set of quantum numbers n, m_s, l, and m by greek indices α, β, . . . , and the wave-functions $u_\alpha(\mathbf{r})$ then form a complete orthonormal basis for the bound electron.

It is tempting to keep this complete orthonormal basis for other atoms as well, and assume that the ground state of an N-electron atom is the tensor product of the N lowest energy eigenstates, appropriately anti-symmetrized via the Slater determinant. Indeed, the periodic table is based on this assumption. However, this ignores the fact that the electrons interact with each other, and the ground state of a many-electron atom is different. The **Hartree-Fock method** is designed to take this into account. It is a constrained variational approach, in which the trial state that is optimised over is forced to be a Slater determinant in order to keep the correct particle statistics. In this section we present the Hartree-Fock method, and arrive at the Hartree-Fock equations, which can be solved iteratively. We follow the derivation given by Bransden and Joachain (Physics of Atoms and Molecules, 1983 pp. 320-339).

First, we specify the Hamiltonian. Using the notation $r_i = |\mathbf{r}_i|$ for the distance of the i^{th} electron from the nucleus, and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ for the distance between electrons i and j, we find that

$$H = H_1 + H_2 = -\sum_{i=1}^{N}\left(\frac{\hbar^2}{2m}\nabla_i^2 + \frac{Z}{r_i}\right) + \sum_{i<j=1}^{N}\frac{1}{r_{ij}}, \tag{9.1}$$

where we used units in which $e/4\pi\epsilon_0 = 1$ and the Hamiltonian is divided into the single electron Hamiltonian (H_1) and the inter-electron Hamiltonian (H_2). We choose as a normalised trial quantum state $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ the Slater determinant

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} u_\alpha(\mathbf{r}_1) & u_\beta(\mathbf{r}_1) & \ldots & u_\upsilon(\mathbf{r}_1) \\ u_\alpha(\mathbf{r}_2) & u_\beta(\mathbf{r}_2) & \ldots & u_\upsilon(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ u_\alpha(\mathbf{r}_N) & u_\beta(\mathbf{r}_N) & \ldots & u_\upsilon(\mathbf{r}_N) \end{vmatrix}, \tag{9.2}$$

and calculate the expectation value of the Hamiltonian H. This must be larger or equal to the ground state $\boldsymbol{E}_0$

$$\langle\Psi|H|\Psi\rangle \geq E_0, \tag{9.3}$$

and varying the trial state then allows us to minimise the expectation value. This will get us close to the ground state energy.

Since the Slater determinant is a rather large expression, it will save us quite a bit of writing if we introduce the anti-symmetrisation operator $\mathscr{A}$, such that

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \sqrt{N!}\mathscr{A}u_\alpha(\mathbf{r}_1)\,u_\beta(\mathbf{r}_2)\ldots u_\upsilon(\mathbf{r}_N) \equiv \sqrt{N!}\mathscr{A}\Phi_{\mathrm{H}}, \tag{9.4}$$

where $\Phi_{\mathrm{H}}(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ is the Hartree wave function. The operator $\mathscr{A}$ can then be written as a sum over permutations P of the labels $\alpha, \beta, \ldots \upsilon$:

$$\mathscr{A} = \frac{1}{N!}\sum_P (-1)^P P, \tag{9.5}$$

and $\mathscr{A}$ is a projection operator: $\mathscr{A}^2 = \mathscr{A} = \mathscr{A}^\dagger$. Both H_1 and H_2 commute with $\mathscr{A}$.

Next, we calculate the expectation values $\langle\Psi|H_1|\Psi\rangle$ and $\langle\Psi|H_2|\Psi\rangle$. Since $[H_1, \mathscr{A}] = 0$, we can write

$$\langle\Psi|H_1|\Psi\rangle = N!\langle\Phi_{\mathrm{H}}|\mathscr{A}H_1\mathscr{A}|\Phi_{\mathrm{H}}\rangle = N!\langle\Phi_{\mathrm{H}}|H_1\mathscr{A}^2|\Phi_{\mathrm{H}}\rangle = N!\langle\Phi_{\mathrm{H}}|H_1\mathscr{A}|\Phi_{\mathrm{H}}\rangle. \tag{9.6}$$

A permutation of the labels $\alpha, \beta, \ldots v$ leads to an orthonormal state and H_1 is a sum over oneelectron Hamiltonians. We can therefore write this as

$$\begin{aligned}\langle \Psi | H_1 | \Psi \rangle &= N! \langle \Phi_{\mathrm{H}} | H_1 \mathscr{A} | \Phi_{\mathrm{H}} \rangle = \sum_P (-1)^P \langle \Phi_{\mathrm{H}} | H_1 P | \Phi_{\mathrm{H}} \rangle = \sum_{i=1}^{N} \left\langle \Phi_{\mathrm{H}} \left| \hat{h}_i \right| \Phi_{\mathrm{H}} \right\rangle \\ &= \sum_{\alpha} \left\langle u_\alpha(\mathbf{r}_i) \left| \hat{h}_i \right| u_\alpha(\mathbf{r}_i) \right\rangle \equiv \sum_{\alpha} I_\alpha\end{aligned} \tag{9.7}$$

where $\hat{h}_i$ is the single electron Hamiltonian

$$\hat{h}_i = -\frac{\hbar^2}{2m}\nabla_i^2 - \frac{Z}{r_i}. \tag{9.8}$$

Next, we calculate the expectation value $\langle \Psi | H_2 | \Psi \rangle$ of the two-electron interaction Hamiltonians. Using the same reasoning as in Eq. (9.6), we find that

$$\langle \Psi | H_2 | \Psi \rangle = N! \langle \Phi_{\mathrm{H}} | H_2 \mathscr{A} | \Phi_{\mathrm{H}} \rangle . \tag{9.9}$$

Substituting the explicit form of $\mathscr{A}$, we find

$$\langle \Psi | H_2 | \Psi \rangle = \sum_{i<j} \sum_P (-1)^P \left\langle \Phi_{\mathrm{H}} \left| \frac{P}{r_{ij}} \right| \Phi_{\mathrm{H}} \right\rangle = \sum_{i<j} \left\langle \Phi_{\mathrm{H}} \left| \frac{1 - P_{ij}}{r_{ij}} \right| \Phi_{\mathrm{H}} \right\rangle , \tag{9.10}$$

where P_{ij} is the exchange operator of electrons i and j. This expression allows us to write

$$\begin{aligned}\langle \Psi | H_2 | \Psi \rangle = \tfrac{1}{2} \sum_{\alpha,\beta} \Big[&\left\langle u_\alpha(\mathbf{r}_i) u_\beta(\mathbf{r}_j) \left| \tfrac{1}{r_{ij}} \right| u_\alpha(\mathbf{r}_i) u_\beta(\mathbf{r}_j) \right\rangle \\ &- \left\langle u_\alpha(\mathbf{r}_i) u_\beta(\mathbf{r}_j) \left| \tfrac{1}{r_{ij}} \right| u_\beta(\mathbf{r}_i) u_\alpha(\mathbf{r}_j) \right\rangle \Big] .\end{aligned} \tag{9.11}$$

Note the swap of α and β in the last ket. This expectation value consists of two terms, namely the direct term

$$J_{\alpha\beta} \equiv \left\langle u_\alpha(\mathbf{r}_i) u_\beta(\mathbf{r}_j) \left| \frac{1}{r_{ij}} \right| u_\alpha(\mathbf{r}_i) u_\beta(\mathbf{r}_j) \right\rangle \tag{9.12}$$

and the exchange term

$$K_{\alpha\beta} \equiv \left\langle u_\alpha(\mathbf{r}_i) u_\beta(\mathbf{r}_j) \left| \frac{1}{r_{ij}} \right| u_\beta(\mathbf{r}_i) u_\alpha(\mathbf{r}_j) \right\rangle \tag{9.13}$$

The total expectation value therefore becomes

$$\langle \Psi | H | \Psi \rangle = \sum_\alpha I_\alpha + \frac{1}{2} \sum_{\alpha,\beta} (J_{\alpha\beta} - K_{\alpha\beta}) . \tag{9.14}$$

The matrix elements $J_{\alpha\beta}$ and $K_{\alpha\beta}$ are real and symmetric in α and β.

The second step towards the Hartree-Fock method is to find the minimum of $E \equiv \langle \Psi | H | \Psi \rangle$ by varying the $u_\alpha(\mathbf{r}_i)$. This means finding $\delta E = 0$. However, we must keep the functions $u_\alpha(\mathbf{r}_i)$ orthonormal to each other, and this imposes N^2 constraints. We can incorporate these constraints in the variational procedure by including Lagrange multipliers $\epsilon_{\alpha\beta}$, and the variational equation becomes

$$\delta E - \sum_{\alpha,\beta} \epsilon_{\alpha\beta} \delta \langle u_\alpha(\mathbf{r}) \mid u_\beta(\mathbf{r}) \rangle = 0. \tag{9.15}$$

There is no explicit reference to electron positions $\mathbf{r}_i$ in $\langle u_\alpha(\mathbf{r}) \mid u_\beta(\mathbf{r}) \rangle$ since we are only interested in its orthonormality properties. The Lagrange multipliers $\epsilon_{\alpha\beta}$ form the elements of a Hermitian matrix.

The variational approach ultimately leads to a set of N coupled equations:

$$\begin{aligned}E_\alpha u_\alpha(\mathbf{r}_i) = \hat{h}_i u_\alpha(\mathbf{r}_i) &+ \sum_\beta \left\langle u_\beta(\mathbf{r}_j) \left| r_{ij}^{-1} \right| u_\beta(\mathbf{r}_j) \right\rangle u_\alpha(\mathbf{r}_i) \\ &- \sum_\beta \left\langle u_\beta(\mathbf{r}_j) \left| r_{ij}^{-1} \right| u_\alpha(\mathbf{r}_j) \right\rangle u_\beta(\mathbf{r}_i) ,\end{aligned} \tag{9.16}$$

which are known as the Hartree-Fock equations. These can be solved by iteration up to any desired precision.

This page titled 9.1: Interacting Electrons in Atomic Shells is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

9.2 Spin Waves in Solids

Consider a system of spins with a nearest-neighbour interaction. For a uniform interaction in all directions, this is described by the Hamiltonian

$$H = \pm J \sum_{(i,j)} \mathbf{S}_i \cdot \mathbf{S}_j = \pm J \sum_{(i,j)} S_{z,i} S_{z,j} + \frac{1}{2}\left(S_{+,i} S_{-,j} + S_{-,i} S_{+,j}\right). \tag{9.17}$$

where $J > 0$ is the coupling strength between the spins, $\sum_{(i,j)}$ is the sum over all neighboring pairs, and $S_\pm = S_x \pm i S_y$. The physics described by this Hamiltonian is known as the Heisenberg model. The sign of the coupling (here made explicit) determines whether the spins want to lign up in parallel $(-J)$ or antiparallel $(+J)$. The former situation describes **ferromagnets**, while the later describes **anti-ferromagnets**. The spin operators for different sites $(i \neq j)$ commute with each other, while the spin operators at the same site $(i = j)$ obey the spin algebra of Eq. (7.24).

Both systems have a well-defined ground state. For the ferromagnet this is the tensor product of the ground state of each individual spin. We are interested in the behaviour of the excitations with respect to this ground state. Due to the large degeneracy in the system (all the spins are of the same species with the same coupling J) the excitations act as identical quasi-particles. Consequently we can describe them using creation and annihilation operators. It turns out that they behave like bosons. You can think of an excitation as a higher spin value at some site that propagates to its neighbors due to the interaction. This is called a spin wave.

Suppose that the spins are aligned in the positive z-direction (so we consider $-J$), and S_z has the maximum eigenvalue s. When the spin is lowered by $\hbar$, this creates an excitation in the system, because the spin is no longer lined up. So the z-component of the spin at site j is given by

$$S_{z,j} = s - \hat{a}_j^\dagger \hat{a}_j, \tag{9.18}$$

where $\hat{a}_j^\dagger \hat{a}_j$ is the operator for the number of excitations at site j. Since $S_\pm$ raise and lower the eigenvalue of S_z, we expect that $S_+ \propto$ and $S_- \propto \hat{a}^\dagger$. When we insist on the commutation relation $[S_+, S_-] = 2S_z$, they become

$$S_{+,j} = \left(2s - \hat{a}_j^\dagger \hat{a}_j\right)^{\frac{1}{2}} \hat{a}_j \quad \text{and} \quad S_{-,j} = \left(2s - \hat{a}_j^\dagger \hat{a}_j\right)^{\frac{1}{2}} \hat{a}_j^\dagger. \tag{9.19}$$

This is known as the Holstein-Primakoff transformation.

For small numbers, the operators $S_\pm$ can be approximated as

$$S_{+,j} \simeq \sqrt{2s}\hat{a}_j \quad \text{and} \quad S_{-,j} \simeq \sqrt{2s}\hat{a}_j^\dagger. \tag{9.20}$$

This allows us to write the Heisenberg Hamiltonian of Eq. (9.17) with $-J$ to lowest order as

$$H = -J \sum_{(i,j)} \left[s^2 + s\left(\hat{a}_i^\dagger \hat{a}_j + \hat{a}_i \hat{a}_j^\dagger - \hat{a}_i^\dagger \hat{a}_i - \hat{a}_j^\dagger \hat{a}_j\right)\right]. \tag{9.21}$$

For a simple cubic lattice of side L, lattice constant a and total number of spins $N = (L/a)^3$ we expect the spin waves to have wave vectors

$$\mathbf{k} = \frac{2\pi}{L}(m, n, o) \quad \text{with} \quad m, n, o \in \mathbb{N}, \tag{9.22}$$

and $1 \leq m, n, o \leq L$. The spin sites must now be indicated by a vector $\mathbf{r}$ instead of a single number j, and the Fourier transformation of $\hat{a}_\mathbf{r}^\dagger$ and $\hat{a}_\mathbf{r}$ is given by

$$\hat{a}_\mathbf{r} = \frac{1}{\sqrt{N}} \sum_\mathbf{k} e^{-i\mathbf{k}\cdot\mathbf{r}} \hat{a}_\mathbf{k} \quad \text{and} \quad \hat{a}_\mathbf{r}^\dagger = \frac{1}{\sqrt{N}} \sum_\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{r}} \hat{a}_\mathbf{k}^\dagger, \tag{9.23}$$

which transforms the Heisenberg Hamiltonian to

$$\begin{aligned} H &= -3Js^2N - \frac{Js}{N}\sum_{\mathbf{r},\mathbf{dk},\mathbf{k}'} e^{i\mathbf{r}(\mathbf{k}-\mathbf{k}')}\left(e^{i\mathbf{d}\cdot\mathbf{q}}-1\right)\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}'} \\ &= -3Js^2N - \frac{Js}{N}\sum_{\mathbf{k}}\epsilon(\mathbf{k})\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}}, \end{aligned} \tag{9.24}$$

where **r** is the position of a lattice site, **d** is the vector from a site to its nearest neighbours, which takes care of the sum over nearest neighbours. This is a diagonal Hamiltonian with eigenenergies

$$\epsilon(\mathbf{k}) = 2J_s\left(3-\cos k_x a-\cos k_y a-\cos k_z a\right). \tag{9.25}$$

This is the dispersion relation for the spin waves, and to lowest order $\left(\cos x \simeq 1-\frac{1}{2}x^2\right)$ it is quadratic:

$$\epsilon(\mathbf{k}) = Jsa^2k^2. \tag{9.26}$$

Spin waves are important when we want to manipulate magnetic properties with high frequency, such as in microwave devices. They carry energy, and are therefore a mechanism for dissipation.

For the case of anti-ferromagnets $(+J)$, the ground state is harder to find. Consider an antiferromagnet that is again a simple cubic lattice with alternating spin $\pm s$ and lattice constant a. We can think of this lattice as two sub-lattices with constant spin, and redefine the spins on the $-s$ sub-lattice according to

$$S_x \to -S_x, \quad S_y \to S_y, \quad \text{and} \quad S_z \to -S_z. \tag{9.27}$$

These operators still obey the commutation relations of spin (which $\mathbf{S}\to-\mathbf{S}$ would not), and the Heisenberg Hamiltonian becomes

$$H = -J\sum_{(i,j)} S_{z,i}S_{z,j} + \frac{1}{2}(S_{+,i}S_{+,j}+S_{-,i}S_{-,j}). \tag{9.28}$$

When we apply the Holstein-Primakoff transformation to this Hamiltonian, to first order we obtain

$$H = -J\sum_{(i,j)}\left[s^2 + s\left(\hat{a}^\dagger_i\hat{a}_i + \hat{b}^\dagger_j\hat{b}_j + \hat{a}_i\hat{b}_j + \hat{a}^\dagger_i\hat{b}^\dagger_j\right)\right], \tag{9.29}$$

where $\hat{a}^\dagger_i$ and $\hat{a}_i$ are the creation and annihilation operators for the spin +s sub-lattice, and $\hat{b}^\dagger_j$ and $\hat{b}_j$ are the creation and annihilation operators for the original spin −s sub-lattice. After the Fourier transform of the creation and annihilation operators we get

$$H = -3Js^2N + 3Js\sum_{\mathbf{k}}\left[\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}} + \hat{b}^\dagger_{-\mathbf{k}}\hat{b}_{-\mathbf{k}} + f(\mathbf{k})\left(\hat{a}_{\mathbf{k}}\hat{b}_{-\mathbf{k}} + \hat{a}^\dagger_{\mathbf{k}}\hat{b}^\dagger_{-\mathbf{k}}\right)\right], \tag{9.30}$$

where $f(\mathbf{k}) = \frac{1}{3}(\cos k_x a + \cos k_y a + \cos k_z a)$.

To find the ground state we must diagonalise H so that it is a sum over number operators. This will involve mixing creation and annihilation operators. This is a unitary transformation that can be written as

$$\hat{a}_{\mathbf{k}} = u_{\mathbf{k}}\hat{c}_{\mathbf{k}} - v_{\mathbf{k}}\hat{d}^\dagger_{-\mathbf{k}} \quad \text{and} \quad \hat{b}_{-\mathbf{k}} = u_{\mathbf{k}}\hat{d}_{-\mathbf{k}} - v_{\mathbf{k}}\hat{c}^\dagger_{\mathbf{k}} \tag{9.31}$$

This leads to the Hamiltonian

$$H = -3Js(s+1)N + \sum_{\mathbf{k}}\epsilon(\mathbf{k})\left(\hat{c}^\dagger_{\mathbf{k}}\hat{c}_{\mathbf{k}} + \hat{d}^\dagger_{-\mathbf{k}}\hat{d}_{-\mathbf{k}} + 1\right), \tag{9.32}$$

with the spin wave energy

$$\epsilon(\mathbf{k}) = 3J_s\left(1-f(\mathbf{k})^2\right)^{\frac{1}{2}} \tag{9.33}$$

For small **k** the dispersion relation of the spin wave is linear in the wave vector, $\epsilon(\mathbf{k}) \simeq \sqrt{3}J_{sak}$, which means that the spin waves behave markedly different in ferromagnets and anti-ferromagnets.

This page titled 9.2 Spin Waves in Solids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

9.3: An Atom in a Cavity

Consider a two-level atom with bare energy eigenstates $|g\rangle$ and $|e\rangle$ and energy splitting $\hbar\omega_0$. The free Hamiltonian H_0 of the atom is given by

$$H_0 = \frac{1}{2}\hbar\omega_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.34}$$

The atom interacts with an electromagnetic wave, and the interaction is approximately the coupling between the dipole moment $\hat{\mathbf{d}}$ associated with the $|g\rangle \leftrightarrow |e\rangle$ transition and the electric field $\mathbf{E}$, leading to the interaction Hamiltonian

$$H_{\text{int}} = -\hat{\mathbf{d}} \cdot \mathbf{E}, \tag{9.35}$$

where $\hat{\mathbf{d}} = -e\hat{\mathbf{r}}$, and $\mathbf{E}$ is the classical, complex-valued electric field at the position of the atom. For an atom at the position $\mathbf{r}$ = 0, the electric field can be written as

$$\mathbf{E} = \mathscr{E}_0 \boldsymbol{\epsilon} e^{i\omega t} + \mathscr{E}_0 \boldsymbol{\epsilon}^* e^{-i\omega t}, \tag{9.36}$$

where $\mathscr{E}_0$ is the real amplitude of the electric field ϵ is the polarsation vector of the wave. The off-diagonal matrix elements of H_{int} are given by

$$\langle e \,|H_{\text{int}}\,| \, g\rangle = e\mathscr{E}_0 \langle e|\hat{\mathbf{r}}|g\rangle \cdot \epsilon e^{i\omega t} + \text{ c.c.}, \tag{9.37}$$

and c.c. denotes the complex conjugate. Since $\hat{\mathbf{r}}$ has odd parity, the diagonal matrix elements $\langle g \,|H_{\text{int}}\,| \, g\rangle$ and $\langle e \,|H_{\text{int}}\,| \, e\rangle$ vanish. When we define $\mathbf{r}_{eg} \equiv \langle e|\hat{\mathbf{r}}|g\rangle$, the total Hamiltonian becomes

$$H = \begin{pmatrix} \frac{1}{2}\hbar\omega_0 & e\mathscr{E}_0 \mathbf{r}_{eg}^* \cdot \left(\epsilon e^{i\omega t} + \epsilon^* e^{-i\omega t}\right) \\ e\mathscr{E}_0 \mathbf{r}_{eg} \cdot \left(\epsilon e^{i\omega t} + \epsilon^* e^{-i\omega t}\right) & -\frac{1}{2}\hbar\omega_0 \end{pmatrix}. \tag{9.38}$$

Using the **Rotating Wave Approximation** (see exercise 9.1), this Hamiltonian can be written as

$$H = \frac{\hbar}{2}\begin{pmatrix} v & \Omega^* \\ \Omega & -v \end{pmatrix}, \tag{9.39}$$

where we made the substitution

$$v = \omega - \omega_0 \quad \text{and} \quad \Omega = \frac{2e\mathscr{E}_0}{\hbar}\mathbf{r}_{eg} \cdot \epsilon. \tag{9.40}$$

We can use the standard matrix techniques in quantum mechanics to solve for the eigenvalues, the eigenstates, and the time evolution of the atom.

Next, we consider the situation where atom is placed inside a cavity of volume V, and the electric field in the cavity has angular frequency ω with wave vector k propagating in the z-direction. Assume that the length of the cavity is a multiple of $\lambda/2$, such that ω is a resonant cavity mode. The field is very weak, so that the classical description of $\mathbf{E}$ is no longer sufficient. In particular, the field is made of photons, i.e., identical bosons. Consequently, we need to express $\mathbf{E}$ in terms of bosonic creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$, which create photons of frequency ω. Since the intensity of the field is proportional to both $\mathscr{E}_0^2$ and $\hat{a}^\dagger\hat{a}$, we expect the electric field to be proportional to the creation and annihilation operators themselves. Hermiticity requires that it is proportional to the sum of the creation and annihilation operators. Furthermore, the electric field is a transverse standing wave cavity mode and must vanish at the mirrors due to the boundary conditions imposed by Maxwell's equations. The spatial amplitude variation therefore includes a factor sin kz. For linear polarisation the operator form of $\mathbf{E}$ then becomes

$$\hat{\mathbf{E}}(z,t) = \epsilon\sqrt{\frac{\hbar\omega}{\epsilon_0 V}}\left(\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t}\right)\sin kz, \tag{9.41}$$

where we assumed that ϵ is real[7]. The creation and annihilation operators are thus the amplitude operators of the field. Note that by the analogy with the harmonic oscillator, the electric field operator acts as a position operator of a particle in a harmonic potential well characterised by ω.

We again consider the dipole approximation of the atom in the field, and the Hamiltonian is written as

$$H_{\text{int}} = -\hat{\mathbf{d}} \cdot \hat{\mathbf{E}} = e\hat{\mathbf{r}} \cdot \epsilon \sqrt{\frac{\hbar\omega}{\epsilon_0 V}} \sin kz \left(\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t} \right). \tag{9.42}$$

The operator $\hat{\mathbf{r}}$ can be written as

$$\hat{\mathbf{r}} = \mathbf{r}_{eg} |e\rangle \langle g| + \mathbf{r}_{eg}^* |g\rangle \langle e|, \tag{9.43}$$

and for notational simplicity, we define the coupling constant g as

$$g = e\mathbf{r}_{eg} \cdot \epsilon \sqrt{\frac{\hbar\omega}{\epsilon_0 V}} \sin kz. \tag{9.44}$$

We again calculate the matrix elements of H_{int} as before, but this time we write the operator in terms of $|g\rangle\langle e|$ and $|e\rangle\langle g|$:

$$H_{\text{int}} = g|e\rangle\langle g| \left(\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t} \right) + g^* |g\rangle\langle e| \left(\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t} \right). \tag{9.45}$$

It is convenient to express $|g\rangle\langle e|$ and $|e\rangle\langle g|$ in terms of the two-level raising and lowering operators σ_+ and σ_-:

$$\sigma_+ = |e\rangle\langle g| \quad \text{and} \quad \sigma_- = |g\rangle\langle e|, \tag{9.46}$$

with the commutation relation

$$[\sigma_+, \sigma_-] = 2\sigma_3 \quad \text{with} \quad \sigma_3 = |g\rangle\langle g| - |e\rangle\langle e| \tag{9.47}$$

The interaction Hamiltonian becomes

$$H_{\text{int}} = g\sigma_+ \left(\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t} \right) + g^* \sigma_- \left(\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t} \right). \tag{9.48}$$

Including the free Hamiltonian for the field and the two-level atom, this becomes in the Rotating Wave Approximation

$$H_{\text{JC}} = \frac{1}{2}\hbar\omega_0\sigma_3 + \hbar\omega\hat{a}^\dagger\hat{a} + g\sigma_+\hat{a} + g^*\sigma_-\hat{a}^\dagger, \tag{9.49}$$

This is the Jaynes-Cummings Hamiltonian for a two-level atom with energy splitting $\hbar\omega_0$ interacting with a cavity mode of frequency ω. To achieve the strongest coupling g, the volume of the cavity should be small, and the atom should sit at an anti-node of the field.

Quantities are conserved when they commute with the Hamiltonian. We can identify two observables that satisfy this requirement, namely the number of electrons

$$\hat{P}_e = |g\rangle\langle g| + |e\rangle\langle e| = \mathbb{I}, \tag{9.50}$$

and the total number of excitations

$$\hat{N}_e = \hat{a}^\dagger\hat{a} + |e\rangle\langle e| \tag{9.51}$$

This means that the Hamiltonian will not couple states with different total excitations.

In a real system, the cavity will not be perfect, and the excited state of the atom will suffer from spontaneous emission into modes other than the cavity mode. This can be modelled by a Lindblad equation for the joint state ρ of the atom and the cavity mode. The Lindblad operator for a leaky cavity is proportional to the annihilation operator $\hat{a}$, with a constant of proportionality $\sqrt{\kappa}$ that denotes the leakage rate. The spontaneous emission of the atom is modelled by the Lindblad operator $\sqrt{\gamma}\sigma_-$. The Lindblad equation then becomes

$$\frac{d\rho}{dt} = \frac{1}{i\hbar}[H_{\text{JC}}, \rho] + \gamma\sigma_-\rho\sigma_+ - \frac{\gamma}{2}\{\sigma_+\sigma_-, \rho\} + \kappa\hat{a}\rho\hat{a}^\dagger - \frac{\kappa}{2}\left\{\hat{a}^\dagger\hat{a}, \rho\right\}. \tag{9.52}$$

The research field of cavity quantum electrodynamics (or cavity QED) is devoted in a large part to solving this equation.

---

[7]This is true for linear polarisation. For elliptical polarisation ϵ will be complex. The subsequent derivation will be slightly modified (with more terms in H_{int}), but no extra technical or conceptual difficulties arise.

This page titled 9.3: An Atom in a Cavity is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

9.4: Outlook: Quantum Field Theory

We have surreptitiously introduced the basic elements of non-relativistic quantum field theory. Consider again the Heisenberg model, where we described a lattice of spins with nearestneighbour interactions. If we take the limit of the lattice constant $a \to 0$ we end up with a continuum of creation and annihilation operators for each point in space. This is a field.

Traditionally we construct a quantum field theory from harmonic oscillators at each point in space. To this end, we characterise a classical harmonic oscillator with mass M in terms of its displacement q and velocity $\dot{q}$. The equations of motion for the classical harmonic oscillator are the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0, \tag{9.53}$$

where L is the Lagrangian

$$L = \frac{1}{2}M\dot{q}^2 - \frac{1}{2}Kq^2, \tag{9.54}$$

and K can be thought of as a spring constant. Substituting this L into Eq. (9.53) yields the familiar differential equations for the harmonic oscillator $\ddot{q} + \Omega^2 q = 0$, with $\Omega^2 = K/M$.

Next, we arrange N particles in a one-dimensional lattice of length L and lattice constant a, where $L = Na$. Each particle's displacement is coupled to the displacement of it's nearest neighbours by a spring with constant K. The equations of motion of this set of coupled particles is given by

$$\ddot{q}_n = \Omega^2 \left[(q_{n+1} - q_n) + (q_{n-1} - q_n)\right]. \tag{9.55}$$

We take the limit of $a \to 0$ and $N \to \infty$ while keeping $L = Na$ fixed. Our variable $q_n(t)$ then becomes a field $u(x,t)$, and it takes only a few lines of algebra to show that

$$\ddot{u}(x,t) = a^2\Omega^2 u''(x,t) = v^2 u''(x,t) \quad \text{with} \quad \lim_{a\to 0} a\Omega = v. \tag{9.56}$$

This is a wave equation, and v is the velocity of the wave. We can generalise this immediately to three dimensions:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} - \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = 0. \tag{9.57}$$

In quantum field theory, we consider only these waves as the field excitations, and ignore the underlying particle structure we used to arrive at this result. We have already done something similar when we considered the spin waves in the Heisenberg model. Note also that the finite speed of propagation of waves means that we can make the theory Lorentz invariant when $v = c$, the speed of light, and Eq. (9.57) becomes $\partial_\mu \partial^\mu u = 0$.

The wave equation is typically derived from a Lagrangian L, or in the case of a field theory, the Lagrangian density $\mathscr{L}$. A massless scalar field is described by the Lorentz-invariant Lagrangian density

$$\mathscr{L} = \frac{1}{2}\left[\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial x}\right)^2 - \frac{1}{c^2}\left(\frac{\partial\phi}{\partial x}\right)^2\right] = \frac{1}{2}\left(\partial_\mu\phi\right)\left(\partial^\mu\phi\right), \tag{9.58}$$

where for technical reasons we redefined $\phi = u/\sqrt{a}$. The dispersion relation for such a field is $c^2k^2 = \omega^2$, with k the wave number and ω the frequency of the wave. Similarly, a massive field is described by

$$\mathscr{L} = \frac{1}{2}\left(\partial_\mu\phi\right)\left(\partial^\mu\phi\right) - \frac{1}{2}m^2\phi^2. \tag{9.59}$$

The Euler-Lagrange equation for this Lagrangian density is the so-called Klein-Gordon equation

$$\left(\partial_\mu\partial^\mu + \frac{m^2c^2}{\hbar^2}\right)\phi = 0. \tag{9.60}$$

The mass term leads to a new dispersion relation

$$c^2\hbar^2k^2 - \hbar^2\omega^2 + m^2c^4 = 0, \tag{9.61}$$

and the group velocity for wave packets is

$$v_g = \frac{d\omega}{dk} = \frac{c}{\sqrt{1+\mu^2}} \quad \text{with} \quad \mu = \frac{mc}{\hbar k}. \tag{9.62}$$

For relativistic particles the momentum $\hbar k$ is much larger than the rest mass mc, and therefore v_g approaches c.

We can solve the Klein-Gordon equation formally by writing

$$\phi(\mathbf{r},t) = \int \frac{d\mathbf{k}}{\sqrt{2(2\pi)^3\omega_\mathbf{k}}} \left(a(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r} - i\omega t} + a^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r} + i\omega t} \right), \tag{9.63}$$

where $a_\mathbf{k}$ is the complex amplitude of a wave with wave vector $\mathbf{k}$. The field is essentially a superposition of (non-interacting) eigenmodes labelled by $\mathbf{k}$, and we call this a free field. We can introduce interactions between the waves in the field by adding higher-order terms to $\mathscr{L}$ with coupling constants $v, \lambda, \ldots$

$$\mathscr{L} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m^2\phi^2 - \frac{v}{3!}\phi^3 - \frac{\lambda}{4!}\phi^4 - \cdots \tag{9.64}$$

The solution to the Klein-Gordon equation is now no longer the correct solution to the new equations of motion, but when the interaction is reasonably weak we can use the solutions $\phi(\mathbf{r},t)$ of the free field as a starting point in a perturbation expansion.

So far, everything in this section has been a classical treatment. In order to extend the theory to quantum mechanics we have to quantise the field. We achieve this by promoting the amplitudes in ϕ (and therefore ϕ itself) to operators that obey commutation of anti-commutation relations)

$$\hat{\phi}(\mathbf{r},t) = \int \frac{d\mathbf{k}}{\sqrt{2(2\pi)^3\omega_\mathbf{k}}} \left(\hat{a}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_\mathbf{k} t} + \hat{a}^\dagger(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r} + i\omega_\mathbf{k} t} \right). \tag{9.65}$$

These are the creation and annihilation operators for excitations of the field. For the Klein-Gordon equation they obey the bosonic commutation relations

$$\left[\hat{a}(\mathbf{k}), \hat{a}\left(\mathbf{k}'\right)\right] = \left[\hat{a}^\dagger(\mathbf{k}), \hat{a}^\dagger\left(\mathbf{k}'\right)\right] = 0 \quad \text{and} \quad \left[\hat{a}(\mathbf{k}), \hat{a}^\dagger\left(\mathbf{k}'\right)\right] = \delta^3\left(\mathbf{k} - \mathbf{k}'\right). \tag{9.66}$$

The state of the field can then be written as a superposition of Fock states. The field $\hat{\phi}$ has now become an observable.

In quantum field theory, the excitations of the field are interpreted as particles. All fundamental particles like quarks, electrons, photons, and the Higgs boson are excitations of a corresponding field. So the excitations of the Higgs field are Higgs bosons, and the excitations of the electromagnetic field are photons. Spin$-\frac{1}{2}$ particles obey the Dirac equation

$$(i\hbar\gamma^\mu\partial_\mu - mc)\,\hat{\psi} = 0, \tag{9.67}$$

where the γ^μ are 4×4 matrices

$$\gamma^0 = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} 0 & \sigma_x \\ -\sigma_x & 0 \end{pmatrix}, \quad \gamma^2 = \begin{pmatrix} 0 & \sigma_y \\ -\sigma_y & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} 0 & \sigma_z \\ -\sigma_z & 0 \end{pmatrix}, \tag{9.68}$$

and ψ is a four-dimensional vector field called a spinor field. The solution to the free Dirac field can be written as

$$\hat{\psi}(\mathbf{r},t) = \sum_s \int \frac{d\mathbf{k}}{\sqrt{2(2\pi)^3\omega_\mathbf{k}}} \left[\hat{b}_s(\mathbf{k}) u_s(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_\mathbf{k} t} + \hat{d}_s^\dagger(\mathbf{k}) v_s(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r} + i\omega_\mathbf{k} t} \right], \tag{9.69}$$

where $s = \pm\frac{1}{2}\hbar$ is the spin value, and $u_s(\mathbf{k})$ and $v_s(\mathbf{k})$ are two spinors carrying the spin component of the field

$$u_s(\mathbf{k}) = \mathscr{N} \begin{pmatrix} \chi_s \\ \frac{\hbar c\sigma\cdot\mathbf{k}}{\hbar\omega + mc^2}\chi_s \end{pmatrix} \quad \text{and} \quad v_s(\mathbf{k}) = \mathcal{N} \begin{pmatrix} -\frac{\hbar c\cdot\sigma\cdot\mathbf{k}}{\hbar\omega + mc^2}\chi_s \\ \chi_s \end{pmatrix}, \tag{9.70}$$

where $\mathscr{N}$ is a normalisation constant, $\boldsymbol{\sigma}$ is a vector of Pauli matrices, and

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{9.71}$$

The creation and annihilation operators $\hat{b}_s(\mathbf{k})$ and $\hat{d}_s^\dagger(\mathbf{k})$ obey anti-commutation relations:

$$\begin{aligned}
&\left\{\hat{b}_s(\mathbf{k}),\hat{b}_r\left(\mathbf{k}'\right)\right\}=\left\{\hat{b}_s^\dagger(\mathbf{k}),\hat{b}_r^\dagger\left(\mathbf{k}'\right)\right\}=0 \quad \text{and} \quad \left\{\hat{b}_s(\mathbf{k}),\hat{b}_r^\dagger\left(\mathbf{k}'\right)\right\}=\delta_{rs}\delta^3\left(\mathbf{k}-\mathbf{k}'\right),\\
&\left\{\hat{d}_s(\mathbf{k}),\hat{d}_r\left(\mathbf{k}'\right)\right\}=\left\{\hat{d}_s^\dagger(\mathbf{k}),\hat{d}_r^\dagger\left(\mathbf{k}'\right)\right\}=0 \quad \text{and} \quad \left\{\hat{d}_s(\mathbf{k}),\hat{d}_r^\dagger\left(\mathbf{k}'\right)\right\}=\delta_{rs}\delta^3\left(\mathbf{k}-\mathbf{k}'\right),\\
&\left\{\hat{b}_s(\mathbf{k}),\hat{d}_r\left(\mathbf{k}'\right)\right\}=\left\{\hat{b}_s^\dagger(\mathbf{k}),\hat{d}_r^\dagger\left(\mathbf{k}'\right)\right\}=0 \quad \text{and} \quad \left\{\hat{b}_s(\mathbf{k}),\hat{d}_r^\dagger\left(\mathbf{k}'\right)\right\}=\left\{\hat{d}_s(\mathbf{k}),\hat{b}_r^\dagger\left(\mathbf{k}'\right)\right\}=0.
\end{aligned} \tag{9.72}$$

This means that the excitations of the Dirac field are fermions with spin $\frac{1}{2}$, such as the electron. You see immediately that $\hat{\psi}$ is not Hermitian due to the appearance of $\hat{d}^\dagger$. This means that $\hat{\psi}$ is not an observable and we cannot think of the Dirac field as a quantised version of a classical observable field. There is no classical analog to the Dirac field. This is a consequence of the fact that the anti-particle of the Dirac excitations are not the same as the particle itself. E.g., the positron is different from the electron. Anti-particles are a quintessentially quantum mechanical phenomenon.

There is of course a lot more to quantum field theory than this. For example, the techniques for doing the perturbation expansion of interacting fields leads to Feynman diagrams, and renormalisation theory must be employed to deal with the infinities that crop up in the perturbation theory. Furthermore, one has to choose the right Lagrangian density, and principles such as gauge invariance and CPT invariance are imposed to constrain the possible choices. This leads ultimately to the extraordinary successful Standard Model of particle physics. It the most fundamental theory of Nature that we have, and it is tested to unprecedented accuracy.

? Exercises

1. The Hamiltonian for a two-level atom in the presence of an electromagnetic wave, as given in Eq. (9.38) depends on the time t. This makes it difficult to solve the Schrödinger equation, so in this exercise we will get rid of the time dependence by applying the Rotating Wave Approximation.
   a. If $|\psi'(t)\rangle = U(t)|\psi(t)\rangle$, find the Schrödinger equation for $|\psi'(t)\rangle$. What is the new Hamiltonian?
   b. Choose $U(t)$ such that
   $$U(t)=\begin{pmatrix} e^{i\omega t/2} & 0 \\ 0 & e^{-i\omega t/2}\end{pmatrix},$$
   and H is given by Eq. (9.38). Show that the new Hamiltonian is given by Eq. (9.39).
2. A two-level atom is placed in a perfect cavity with an electromagnetic field of frequency ω.
   a. Show that the Jaynes-Cummings Hamiltonian can be written as a direct sum of 2×2 matrices H_n, and specify H_n.
   b. Diagonalize H_n to find the energy values of the system, and calculate the eigenstates.
   c. At $t=0$, the system is in the state $|\psi(0)\rangle = |e,n\rangle$. Calculate the state $|\psi(t)\rangle$ when the light is on resonance with the atomic transition $(\omega=\omega_0)$.
   d. Calculate the amount of entanglement between the atom and the cavity field. Use the relative entropy as the entanglement measure.
   e. How long does the atom need to reside in the cavity in order to achieve maximum entanglement?

This page titled 9.4: Outlook: Quantum Field Theory is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

10: Pauli Spin Matrices

We can represent the eigenstates for angular momentum of a spin-1/2 particle along each of the three spatial axes with column vectors:

$$|+z\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad |+y\rangle = \begin{bmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{bmatrix} \quad |+x\rangle = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$
$$|-z\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad |-y\rangle = \begin{bmatrix} i/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad |-x\rangle = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

Similarly, we can use matrices to represent the various spin operators.

10.1: Spin Operators
10.2: Expectation Values
10.3: Total Angular Momentum

This page titled 10: Pauli Spin Matrices is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

10.1: Spin Operators

We've been talking about three different spin observables for a spin-1/2 particle: the component of angular momentum along, respectively, the x, y, and z axes. In quantum mechanics, there is an operator that corresponds to each observable. The operators for the three components of spin are $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$. If we use the column vector representation of the various spin eigenstates above, then we can use the following representation for the spin operators:

$$\hat{S}_x = \frac{\hbar}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \hat{S}_y = \frac{\hbar}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \hat{S}_z = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{10.2}$$

It is also conventional to define the three "Pauli spin matrices" σ_x, σ_y, and σ_z, which are:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{10.3}$$

Clearly, then, the spin operators can be built from the corresponding Pauli matrices just by multiplying each one by $\hbar/2$.

You can verify that this is a good representation of the spin operators by making sure that all all of the various observations about spin states are reproduced by using these operators and these vectors to predict them from the theory. For example, $|+y\rangle$ is an eigenstate for the y component of spin, so the column vector representation of $|+y\rangle$ needs to be an eigenvector of $\hat{S}_y$. Is it? Let's try it:

$$\begin{aligned} \hat{S}_y|+y\rangle &= \frac{\hbar}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{bmatrix} \\ &= \frac{\hbar}{2}\begin{bmatrix} (0)(1/\sqrt{2})+(-i)(i/\sqrt{2}) \\ (i)(1/\sqrt{2})+(0)(i/\sqrt{2}) \end{bmatrix} \\ &= \frac{\hbar}{2}\begin{bmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{bmatrix} \\ &= \frac{\hbar}{2}|+y\rangle \end{aligned} \tag{10.4}$$

In at least this case, the matrix and column vector representations of $\hat{S}_y$ and $|+y\rangle$ are working.

This page titled 10.1: Spin Operators is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

10.2: Expectation Values

You can also use the matrix representation of operators to figure out expectation values. Suppose that you have an electron in the state:

$$|\psi\rangle = \sqrt{\frac{1}{3}}|+z\rangle + \sqrt{\frac{2}{3}}|-z\rangle \tag{10.5}$$

What are the expectation values of for spin along the x-axis??

First, we construct the column vector representation of this state $|\psi\rangle$:

$$|\psi\rangle = \begin{bmatrix} \sqrt{1/3} \\ \sqrt{2/3} \end{bmatrix} \tag{10.6}$$

The corresponding bra vector is represented by a row vector:

$$\langle\psi| = \begin{bmatrix} \sqrt{1/3} & \sqrt{2/3} \end{bmatrix} \tag{10.7}$$

To figure out the expectation value of x-spin, we sandwich the $\hat{S}_x$ operator in between the bra and ket vectors for this state:

$$\begin{aligned} \langle s_x \rangle &= \left\langle \psi \middle| \hat{S}_x \middle| \psi \right\rangle \\ &= \begin{bmatrix} \sqrt{1/3} & \sqrt{2/3} \end{bmatrix} \left(\frac{\hbar}{2}\right) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{1/3} \\ \sqrt{2/3} \end{bmatrix} \\ &= \left(\frac{\hbar}{2}\right) \begin{bmatrix} \sqrt{1/3} & \sqrt{2/3} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{1/3} \\ \sqrt{2/3} \end{bmatrix} \end{aligned} \tag{10.8}$$

(all we did between the last two lines was pull the scalar constant $\hbar/2$ out front). We've got a row vector times a matrix times a column vector. That may look intimidating, but we know how to do the matrix times the column vector, so let's do that first. That will leave us with a row vector times a column vector; we know how to work that out as well, leaving us with just a scalar. A scalar is what we need for an expectation value.

$$\begin{aligned} \langle s_x \rangle &= \left(\frac{\hbar}{2}\right) \begin{bmatrix} \sqrt{1/3}\sqrt{2/3} \end{bmatrix} \begin{bmatrix} \sqrt{2/3} \\ \sqrt{1/3} \end{bmatrix} \\ &= \left(\frac{\hbar}{2}\right) \left(\sqrt{\frac{1}{3}}\right) \left(\sqrt{\frac{2}{3}}\right) + \left(\sqrt{\frac{2}{3}}\right) \left(\sqrt{\frac{1}{3}}\right) \\ &= \left(\frac{\hbar}{2}\right) \frac{2\sqrt{2}}{3} \\ &= \frac{\sqrt{2}}{3}\hbar \end{aligned} \tag{10.9}$$

That's a plausible expectation value. It's neither $\hbar/2$ nor $-\hbar/2$, which means that this is not a definite state for x spin. That's good, because the state is clearly not the same as $|+x\rangle$ when you write out that state in terms of $|+z\rangle$ and $|-z\rangle$. It's between those two. However, from just looking at the state, while you can fairly quickly see that $|-z\rangle$ has more amplitude than $|+z\rangle$, and thus a measurement of z spin will yield $-\hbar/2$ more often than $+\hbar/2$, it's not obvious at all just looking at the state which value of x spin would be more common, and thus whether the x expectation value should be positive or negative. In this case, you have to perform the calculation. The matrix formulation of the spin operators makes the calculations faster and easier than they would be when you explicit writing out everything in terms of the z basis states.

We could also quickly figure out what the amplitude for measuring positive x spin is with this formalism. Remember that for a particle in state $|\psi\rangle$, the amplitude for finding positive x spin is $\langle +x \mid \psi\rangle$. Putting together the $\langle +x|$ bra vector with the column vector for $|\psi\rangle$ above, we get:

$$\begin{aligned}\langle +x \mid \psi \rangle &= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{1/3} \\ \sqrt{2/3} \end{bmatrix} \\ &= \left(\sqrt{\frac{1}{2}}\right)\left(\sqrt{\frac{1}{3}}\right) + \left(\sqrt{\frac{1}{2}}\right)\left(\sqrt{\frac{2}{3}}\right) \\ &= \sqrt{\frac{1}{6}} + \sqrt{\frac{1}{3}} \\ &= 0.9856\end{aligned} \tag{10.10}$$

That's a high positive amplitude, corresponding to a probability of 0.97 that positive x spin will be measured for this state. Again, without performing the calculations, this is not at all obvious. However, this high probability for positive x spin is consistent with the fact that the x spin expectation value $\langle s_x \rangle$ is positive and only a little bit less than $\hbar/2$.

This page titled 10.2: Expectation Values is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

10.3: Total Angular Momentum

In 3D space, if you have three components of a vector $\vec{v}$, then the magnitude of that vector squared is $v^2 = v_x^2 + v_y^2 + v_z^2$. Angular momentum is a vector, and so this rule would apply to angular momentum as well. However, in quantum mechanics, we see that angular momentum behaves very differently from how it does in classical physics. In particular, if an object has a definite z component of angular momentum, then it has an indefinite x component of angular momentum. Does that mean that total angular momentum must also be indefinite? In order to answer this question, we must ask it in a proper quantum manner.

In quantum mechanics, we associate each observable quantity with an operator. We can then use that operator on one of its eigenstates (i.e. a state where the observable has a definite value) to pull out the value of the observable as the eigenvalue a in the equation $\hat{A}|\phi\rangle = a|\phi\rangle$. If the system is not in an eigenstate, we can figure out the "expectation value" $\langle a\rangle$ (i.e. the weighted average of all values that could be observed) using the operator in the equation

$$\langle a\rangle = \langle\psi|\hat{A}|\psi\rangle \tag{10.11}$$

What then is the operator that corresponds to total angular momentum? By analogy to classical physics, we can guess that the operator for total angular momentum squared is:

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 \tag{10.12}$$

This begs the question as to what $\hat{S}_x^2$ means; we know how to square numbers and variables, but how do you square an operator? For $\hat{S}^2$, we will just treat that as an operator itself, the "total spin squared" operator. We'll not treat the superscript 2 as squaring, rather we'll just consider it part of the name. As for the others, what really matters about an operator is what it does to a vector representing a quantum state that it's supposed to operate on. In order to make this definition of the spin angular momentum squared operator to work, we need to interpret them as follows:

$$\hat{S}_x^2|\psi\rangle = \hat{S}_x\hat{S}_x|\psi\rangle \tag{10.13}$$

In other words, first apply the $\hat{S}_x$ operator to the state $|\psi\rangle$, and then apply the $\hat{S}_x$ operator again to the vector that resulted from the first application. As an example, let's consider the $\hat{S}_x^2$ operator on the state $|+z\rangle$:

$$\begin{aligned}\hat{S}_x^2|+z\rangle &= \frac{\hbar}{2}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\frac{\hbar}{2}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}\\ &= \frac{\hbar^2}{4}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}\end{aligned} \tag{10.14}$$

All we've done in the first step is pulled the scalar constants out front. To perform this matrix multiplication, first we must multiply the rightmost matrix by the vector, and then we can multiply the first matrix by the result.[1]

$$\begin{aligned}\hat{S}_x^2|+z\rangle &= \frac{\hbar^2}{4}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix}\\ &= \frac{\hbar^2}{4}\begin{bmatrix}1\\ 0\end{bmatrix}\\ &= \frac{\hbar^2}{4}|+z\rangle\end{aligned} \tag{10.15}$$

Interestingly, it seems that $|+z\rangle$ is in fact an eigenstate of $\hat{S}_x^2$, even though it's not an eigenstate of $\hat{S}_x$!

Armed with these techniques, it is possible to show that any properly normalized spin-1/2 state $|\psi\rangle$ is an eigenstate of $\hat{S}_2$ with eigenvalue $\frac{3}{4}\hbar^2$. Although it may be surprising that $|+z\rangle$ is an eigenstate of $\hat{S}_x^2$, in retrospect it should not be surprising that all states are eigenstates of the total angular momentum operator. We've been saying all along that the total angular momentum of an electron is $\frac{\sqrt{3}}{2}\hbar$; what can be in an indefinite state is the components of that angular momentum along various axes.

[1]If you know how to multiply 2×2 matrices, you can do the matrix multiplication first if you wish. As we will see, the commutative property does not apply to matrix multiplication, but the associative property does.

This page titled 10.3: Total Angular Momentum is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

11: Noncommuting Operators and Uncertainty

11.1: Eigenstates and Commuting Operators
11.2: Non-Commuting Operators
11.3: Quantifying Uncertainty
11.3.1: Mean and Variance
11.3.2: Uncertainty in Quantum Mechanics
11.4: The Heisenberg Uncertainty Principle
11.4.1: An Alternate Formulation
11.4.2: Vacuum Energy

This page titled 11: Noncommuting Operators and Uncertainty is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

11.1: Eigenstates and Commuting Operators

You are probably used to the idea that multiplication is commutative. That is, if you have a product ab where a and b are scalars, you can write the multiplication in either order (ab or ba), and the product is exactly the same. This is not necessarily the case for matrix multiplication! If A and B are matrices, then $AB \neq BA$ in general. Sometimes it will be true, but not always. Because we can use matrices to represent operators in quantum mechanics, this means that operators don't commute in general. That is, for example, $\hat{S}_x \hat{S}_y |\psi\rangle \neq \hat{S}_y \hat{S}_x |\psi\rangle$.

Sometimes, however, operators do commute. Suppose that you have two observables A and B with corresponding operators $\hat{A}$ and $\hat{B}$. Suppose also that you have a state $|\phi\rangle$ that is a definite state for both A and B. That means that in our mathematical formalism, $|\phi\rangle$ must be an eigenvector for both $\hat{A}$ and $\hat{B}$:

$$\begin{aligned} \hat{A}|\phi\rangle &= a|\phi\rangle \\ \hat{B}|\phi\rangle &= b|\phi\rangle \end{aligned} \tag{11.1}$$

Here, a and b are the eigenvalues for $\hat{A}$ and $\hat{B}$ respectively. In other words, $|\phi\rangle$ has a definite value of observable A, and a is that value; likewise, it has a definite value of observable B, and b is that value.

Let us now consider the application of both of these operators to this state $|\phi\rangle$:

$$\begin{aligned} \hat{A}\hat{B}|\phi\rangle &= \hat{A}b|\phi\rangle \\ &= b\hat{A}|\phi\rangle \\ &= ba|\phi\rangle \\ &= ab|\phi\rangle \end{aligned} \tag{11.2}$$

where in the last step we've used the fact that a and b are real numbers, so the product of the two of them does in fact commute. Let's now try this in the other order:

$$\begin{aligned} \hat{B}\hat{A}|\phi\rangle &= \hat{B}a|\phi\rangle \\ &= a\hat{B}|\phi\rangle \\ &= ab|\phi\rangle \end{aligned} \tag{11.3}$$

Here, we can see that in fact the operators $\hat{A}$ and $\hat{B}$ do commute if they are operating on a state that is an eigenstate for both operators.

Remember that in the case of spin, we argued that $|+z\rangle$ and $|-z\rangle$ form a complete basis set of vectors; that is, any spin state $|\psi\rangle$ can be written as a sum of scalar constants times those two vectors. In general, a complete set of eigenvectors for a given operator do form a basis set that can be used to construct any vector that is part of the overall scheme that that operator is part of. (For instance, the projection of spin along all three axes are part of the same scheme, as they are all the same kinds of states— that is, spin angular momentum states.) Therefore, we can write any state $|\psi\rangle$ as a sum of constants times the eigenvectors for that operator. If $\hat{A}$ and $\hat{B}$ are two operators that share the same eigenvectors, then $\hat{A}\hat{B}|\psi\rangle = \hat{B}\hat{A}|\psi\rangle$. That is, the operation of these two operators on any state commutes. For that reason, we generally just say that the operators commute.

This page titled 11.1: Eigenstates and Commuting Operators is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

11.2: Non-Commuting Operators

In the previous section, we saw that if a particle can be in a definite state for two observables, then the two operators associated with those observables will commute. The converse is therefore also true; if two operators do not commute, then it is not possible for a quantum state to have a definite value of the corresponding two observables at the same time.

We've already seen examples of this. A particle can't have a definite x spin and a definite y spin at the same time. If our theory is to be useful, then we would hope that $\hat{S}_x$ and $\hat{S}_y$ would not commute when they operate on a general normalized state $|\psi\rangle$. Let's try it first in one order:

$$\begin{aligned}\hat{S}_x\hat{S}_y|\psi\rangle &= \frac{\hbar^2}{4}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix}\\ &= \frac{\hbar^2}{4}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}-i\psi_2\\ i\psi_1\end{bmatrix}\\ &= \frac{\hbar^2}{4}\begin{bmatrix}i\psi_1\\ -i\psi_2\end{bmatrix}\\ &= i\frac{\hbar^2}{4}\begin{bmatrix}\psi_1\\ -\psi_2\end{bmatrix}\end{aligned} \tag{11.4}$$

Now let's try it in the other order:

$$\begin{aligned}\hat{S}_y\hat{S}_x|\psi\rangle &= \frac{\hbar^2}{4}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix}\\ &= \frac{\hbar^2}{4}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\begin{bmatrix}\psi_2\\ \psi_1\end{bmatrix}\\ &= \frac{\hbar^2}{4}\begin{bmatrix}-i\psi_1\\ i\psi_2\end{bmatrix}\\ &= -i\frac{\hbar^2}{4}\begin{bmatrix}\psi_1\\ -\psi_2\end{bmatrix}\end{aligned} \tag{11.5}$$

Clearly the two are not the same; one is the negative of the other. Therefore, $\hat{S}_x$ and $\hat{S}_y$ do not commute when operating on a general state ψ, as expected.

It is interesting to note the effect of $\hat{S}_z$ on this same general state:

$$\begin{aligned}\hat{S}_z|\psi\rangle &= \frac{\hbar}{2}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix}\\ &= \frac{\hbar}{2}\begin{bmatrix}\psi_1\\ -\psi_2\end{bmatrix}\end{aligned} \tag{11.6}$$

Notice that except for the constant out front, the vector produced by $\hat{S}_z$ on this state is the same as the vector produced by $\hat{S}_x\hat{S}_y$ and $\hat{S}_y\hat{S}_x$. In fact, we can put the two together:

$$\begin{aligned}\left(\hat{S}_x\hat{S}_y - \hat{S}_y\hat{S}_x\right)|\psi\rangle &= i\frac{\hbar^2}{2}|\psi\rangle\\ \left[\hat{S}_x, \hat{S}_y\right]|\psi\rangle &= i\hbar\hat{S}_z|\psi\rangle\end{aligned} \tag{11.7}$$

The term in brackets, $\left[\hat{S}_x, \hat{S}_y\right]$ is called the commutator of $\hat{S}_x$ and $\hat{S}_y$. It's defined by the term in parentheses above it: $\left(\hat{S}_x\hat{S}_y - \hat{S}_y\hat{S}_x\right)$. It works out for the commutators of all three spin angular momentum operators that:

$$\left[\hat{S}_x, \hat{S}_y\right] = i\hbar\hat{S}_z \tag{11.8}$$

$$\left[\hat{S}_y, \hat{S}_z\right] = i\hbar\hat{S}_x$$
$$\left[\hat{S}_z, \hat{S}_x\right] = i\hbar\hat{S}_y \tag{11.9}$$

This page titled 11.2: Non-Commuting Operators is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

11.3: Quantifying Uncertainty

If a system is in an indeterminate state for a given observable, it means that we can't know exactly what we're going to measure if we do in fact make a measurement of that observable. We've seen that we can calculate the average of all the measurements we might make, suitably weighted by their probabilities; that's what is called the "expectation value" in quantum mechanics. And, we have seen how we can calculate the amplitude, and from that the probability, that we'll get any given possible measurement. With the spin-1/2 system we've been talking about, there are only two possible values that we might measure for the spin along any given axis: $+\hbar/2$ and $-\hbar/2$. As such, it's convenient just to list the probability of each. If you talk about other systems where there are a lot of possible measurements (including continuous systems such as the position of a particle), it becomes impractical to list the probabilities of each state. It would be nice to have some other way of quantifying our uncertainty.

This page titled 11.3: Quantifying Uncertainty is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

11.3.1: Mean and Variance

Suppose you have a set of values a_j. By saying that this is a set, we mean that we have several values a_1, a_2, a_3, and so forth. The notation a_j, in this context, means that j can be replaced by any integer between 1 and the total number of values that you have in order to refer to that specific value. Suppose that we have N total values. The average of all of our values can be written as:

$$\langle a \rangle = \frac{1}{N} \sum_j a_j \tag{11.10}$$

The letter Σ is the capital Greek letter "sigma". This notation means that you sum together all of the values of a_j that you have. For instance, suppose you had just four values, a_1, a_2, a_3, and a_4, then:

$$\sum_j a_j = a_1 + a_2 + a_3 + a_4 \tag{11.11}$$

Therefore, the mean (or average) value of a in this context is:

$$\langle a \rangle = \frac{1}{N} \sum_j a_j = \frac{1}{N}(a_1 + a_2 + a_3 + a_4) \tag{11.12}$$

To quantify the uncertainty on a set of values, we want to say something about how far, on average, a given value is from the mean of all the values. Thus, it's tempting to try to define the uncertainty as follows:

$$\frac{1}{N} \sum_j (a_j - \langle a \rangle) \tag{11.13}$$

Remember that addition is commutative. Realizing that the $\sum$ symbol just indicates a sum, i.e. a whole lot of addition, we can rewrite this as:

$$\frac{1}{N} \left(\sum_j a_j - \sum_j \langle a \rangle \right) \tag{11.14}$$

The second term in the subtraction is a sum over j of the average value. The average value doesn't depend on which a_j we're talking about; it's a constant, it's the same for all of them. Therefore, the sum of that number N times is just going to be equal to $N\langle a \rangle$. Making this substitution and distributing the $1/N$ into the parentheses:

$$\frac{1}{N} \sum_j a_j - \frac{1}{N} N \langle a \rangle \tag{11.15}$$

But we recognize the first term in this subtraction as just $\langle a \rangle$. So, the total result of this is zero. Clearly, this is not a good expression for the uncertainty in a. If you think about it, the average deviation of a_j from $\langle a \rangle$ ought to be zero. If $\langle a \rangle$ is the average value of a, then a_j should be below $\langle a \rangle$ about as often as it is above, so your sum will have a mix of positive and negative terms. The very definition of the average insures that this sum will be zero.

Instead, we shall define the variance as:

$$\Delta a^2 = \frac{1}{N} \sum_j \left(a_j - \langle a \rangle\right)^2 \tag{11.16}$$

Here, we're using Δa to indicate the uncertainty in a. The variance is defined as the uncertainty squared.[1] The advantage of this expression is that because we're squaring the difference between each value a_j and the average value, we're always going to be summing together positive terms; there will be no negative terms to cancel out the positive terms. Therefore, this should be a reasonable estimate of how far, typically, the measurements a_j are from their average.

We can unpack this sum a bit, first by multiplying out the squared polynomial:

$$\Delta^2 = \frac{1}{N} \sum_j \left(a_j^2 - 2\langle a \rangle a_j + \langle a \rangle^2 \right) \tag{11.17}$$

In order to clean this expression up, inside the parentheses both add and subtract $\langle a \rangle^2$:

$$\begin{aligned}\Delta a^2 &= \frac{1}{N}\sum_j \left(a_j^2 - 2\langle a\rangle a_j + 2\langle a\rangle^2 - \langle a\rangle^2\right) \\ &= \frac{1}{N}\sum_j \left(a_j^2 - \langle a\rangle^2 + 2\langle a\rangle\left(\langle a\rangle - a_j\right)\right) \\ &= \frac{1}{N}\sum_j a_j^2 - \frac{1}{N}\sum_j \langle a\rangle^2 + \frac{1}{N}2\langle a\rangle\sum_j\left(\langle a\rangle - a_j\right)\end{aligned} \tag{11.18}$$

Notice that the last term is going to be zero, as it includes the average difference between the mean and each observation. The second term is just going to be $\langle a\rangle^2$, because once again $\langle a\rangle$ is the same for all terms of the sum; the sum will yield $N\langle a\rangle^2$, canceling the N in the denominator. So, we have:

$$\Delta a^2 = \left\langle a^2\right\rangle - \langle a\rangle^2 \tag{11.19}$$

[1]If you know statistics, you may recognizing this as being very similar to how variance is defined there— only in statistics, we divide by $N-1$ rather than by N. The difference becomes unimportant as N gets large.

This page titled 11.3.1: Mean and Variance is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

11.3.2: Uncertainty in Quantum Mechanics

In order to bring this into quantum mechanics, we already know how to calculate the average $\langle a \rangle$, which we call the "expectation value". If the state of the system is $|\psi\rangle$ and the operator corresponding to the observable a is $\hat{A}$, then

$$\langle a \rangle = \langle \psi | \hat{A} | \psi \rangle \tag{11.20}$$

Similarly, now that we recognize that we can interpret $\hat{A}^2$ as just applying the operator $\hat{A}$ twice, we can calculate $\langle a^2 \rangle$:

$$\langle a^2 \rangle = \left\langle \psi \left| \hat{A}^2 \right| \psi \right\rangle \tag{11.21}$$

For example, let's consider the state $|\psi\rangle = |+z\rangle$ and the observable spin-z. We expect the uncertainty here to be zero, because we know exactly what we'll get if we measure spin-z. Let's see if it works out that way:

$$\begin{aligned} \langle s_z \rangle &= \left\langle \psi \left| \hat{S}_z \right| \psi \right\rangle \\ &= \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{\hbar}{2} \end{aligned} \tag{11.22}$$

As expected, the expectation value for spin-z is $+\hbar/2$. For the other part:

$$\begin{aligned} \langle s_z^2 \rangle &= \left\langle +z \left| \hat{S}_z \hat{S}_z \right| -z \right\rangle \\ &= \frac{\hbar^2}{4} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{\hbar^2}{4} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{\hbar^2}{4} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{\hbar^2}{4} \end{aligned} \tag{11.23}$$

If we take the difference $\langle {s_z}^2 \rangle - \langle s_z \rangle^2$, we get $\hbar^2/4 - \hbar^2/4 = 0$, as expected.

What if we want to know the uncertainty on S_x for this state?

$$\begin{aligned} \langle s_x \rangle &= \left\langle +z \left| \hat{S}_x \right| +z \right\rangle \\ &= \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= 0 \end{aligned} \tag{11.24}$$

If the system is in the state $|+z\rangle$, we know that we have a 50% chance each for finding spin-x to be $+\hbar/2$ or $-\hbar/2$. Thus, it's no surprise that the average value of spin-x is zero, even though zero isn't a value we might measure. To figure out the variance:

$$\begin{aligned}\langle s_x^2\rangle &= \left\langle +z\middle|\hat{S}_x\hat{S}_x\middle|+z\right\rangle \\ &= \frac{\hbar^2}{4}\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix} \\ &= \frac{\hbar^2}{4}\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix} \\ &= \frac{\hbar^2}{4}\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix} \\ &= \frac{\hbar^2}{4}\end{aligned} \tag{11.25}$$

Thus, in this case, the formal uncertainty Δs_x on the x-spin is $\hbar/2$.

This page titled 11.3.2: Uncertainty in Quantum Mechanics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

11.4: The Heisenberg Uncertainty Principle

As previously stated, quantifying the uncertainty on a given observable for a given quantum state is more interesting when the observable we're talking about has a large number (or even a continuum) of different values it might take on. If you consider two different observables whose operators do not commute, then a system cannot be in a definite state for both of those observables at the same time. The Heisenberg Uncertainty Principle takes this observation, makes it stronger, and quantifies it.

Consider a quantum particle that can move along one direction. Its position is then x, and its momentum along that direction is p_x. The Heisenberg Uncertainty Principle states that:

$$\Delta x \Delta p_x \geq \frac{\hbar}{2} \tag{11.26}$$

We've gotten used to thinking of $\hbar$ as an angular momentum unit, because that's where it's shown up before. However, here, it's not really an angular momentum, though it still does of course have the same units (position times position over time). Instead, it represents the fundamental limit in quantum mechanics on how well you can know two different observables, position and momentum. If you know one of them perfectly, e.g. $\Delta x = 0$, then the uncertainty in the other one must be infinite. Although in more advanced quantum mechanics we use such states as they're a good approximation for a lot of things, they're not really physical. In reality, most quantum systems have a small amount of uncertainty in both position and momentum. That is, a particle doesn't have a definite position or a definite momentum, but the range of positions for which it has an appreciable amplitude is confined to a small space, and the range of momenta for which it has an appreciable amplitude is confined to a small range.

As a concrete example, let's consider an electron. For a non-relativistic electron, its momentum is just $p = mv$, where m is its mass and v is its speed. Therefore, $\Delta p = m\Delta v$, as the mass is well known and there is no uncertainty in it. What is a good uncertainty in speed to consider? For practical purposes, let's suppose that we're doing an experiment with an electron that requires it to be localized for 1 second. We don't want the uncertainty in the speed of the electron to cause our uncertainty in the position after one second to be greater than the uncertainty in the position was in the first place. So, we shall choose $\Delta v = \Delta x/t$, where we'll put in $t = 1$ second. If we then put this into the uncertainty principle

$$\begin{aligned} \Delta x \Delta p &\geq \frac{\hbar}{2} \\ \Delta x m_e \frac{\Delta x}{t} &\geq \frac{\hbar}{2} \\ \Delta x &\geq \sqrt{\frac{\hbar t}{2m_e}} \end{aligned} \tag{11.27}$$

If you put in the numbers, you find that the uncertainty on the position of this electron is 0.01 m, or one centimeter. For an electron, that's a lot! (One could argue about whether or not 1 second is a reasonable timescale. When we get to talking about atoms, we'll think more carefully about what a reasonable timescale is.)

Notice, however, that the uncertainty in the position goes down as the mass goes up. Imagine that you stood still your entire life. If you want to balance the uncertainty in your starting position with the uncertainty in your position resulting from the uncertainty in your velocity over your entire life, then you'd put in your age for t. Let us assume, optimistically, that you will live 100 years (3×10^9 seconds), and that your mass is 80 kg. If you put those numbers in to the equation above, you find out that the uncertainty on your position is 4×10^{-14} m. In other words, even though quantum uncertainty can be pretty important for an electron, on everyday scales for macroscopic objects the effect of quantum uncertainty is utterly negligible.

This page titled 11.4: The Heisenberg Uncertainty Principle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

11.4.1: An Alternate Formulation

An alternate way to formulate Heisenberg's Uncertainty principle is:

$$\Delta E \Delta t \geq \frac{\hbar}{2} \tag{11.28}$$

The interpretation of this is a little less clear than in the case of position and momentum. Uncertainty in energy seems obvious enough; it's the square root of the variance of all the values of energy that might be measured for a particle in a given quantum state. But what is "uncertainty on time"? Rather than interpreting this as an uncertainty, we shall interpret it as a time interval. In a sense, that's the same thing; an interval of time is qualitatively similar to an uncertainty on what time it is.

What this means, then, is that the uncertainty in the energy of a quantum state is related to how long that state hangs around. If a system is in an energy eigenstate, then it has a definite energy and $\Delta E = 0$. Such a state must be stable then, for Δt has to be infinite. In other words, in the absence of any interactions, a particle in an energy eigenstate will stay, forever, in an energy eigenstate.

For small time intervals, however, there will be a finite uncertainty in the energy of a system. One thing that this means is that it becomes possible to violate the conservation of energy, so long as you do it so fast that nobody can catch you at it! Among other things, this leads to the possibility of quantum tunneling– that is, if a particle is up against a potential barrier it doesn't have enough energy to penetrate, there is some finite probability that the particle may be located inside the barrier. And, the particle may be able to cross the barrier, even though classically it could not.

Later, when we talk about atoms, states other than the ground state (i.e. lowest energy state) of the atom aren't going to be perfectly stable. Over time, they will decay to the ground state, with a characteristic lifetime analogous to the half-life of a radioactive isotope. Although we will describe these excited states as being energy eigenstates, the fact that they decay tells us that they can't exactly be energy eigenstates. It also tells us that there must be some uncertainty as to the exact energy value associated with those states. There will be observational consequences of this, although in practice for real atoms these consequences are extremely difficult to observe.

This page titled 11.4.1: An Alternate Formulation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

11.4.2: Vacuum Energy

A second consequence of this formulation of the Heisenberg Uncertainty Principle is the possibility of vacuum energy. Consider a small region of space. Suppose that it's empty; that is, you've taken out everything you can take out of it, including atoms, light (photons), dark matter, and so forth. Make sure that there are no quantum systems anywhere with non-negligible probability for being found in this region of space. Over a finite time interval Δt, you can't be sure exactly how much energy there is in this region of space; your uncertainty in the amount of energy must be at least $\Delta E = \frac{\hbar}{2\Delta t}$. As a result, there may be energy in the vacuum.

What is the expectation value of this energy? You might predict that the expectation should be 0, even though the uncertainty has to be greater than zero. Figuring it out requires going into relativistic quantum mechanics, called quantum field theory. Unfortunately, even quantum field theory can't calculate that right, for naive estimates of what you'd get (the best we can really do) gives a value of the vacuum energy density that is so high that it would prevent galaxies from ever having formed in our Universe. The fact that you are reading this indicates that this estimate cannot be right. Indeed, quantum field theory estimates a value for the vacuum energy density that is 120 orders of magnitude too big! That's pretty far off. As such, we have to say that we don't completely understand the nature of vacuum energy.

What form would this vacuum energy take? We've already seen that in a finite time interval Δt, we can't say with certainty that the vacuum has zero energy. In quantum field theory, it becomes possible to create and destroy particles, as long as you obey all of the conservation laws. For example, two photons can interact and create an electron/positron pair, where a positron is the antimatter partner to an electron. If you don't have to worry about conserving energy, however, you can create a positron/electron pair out of absolutely nothing. . . as long as they re-annihilate back to absolutely nothing fast enough. For every fundamental particle that exists, this sort of thing is going on around us all the time.

What is the net energy density of the vacuum as a result of all of this? For a long time, many physicists assumed that a various terms would cancel out to zero; the naive calculations indicated something absurd, and the most natural result if those calculations are wrong is that things would cancel out. However, in the last ten years, observations of the expansion of the Universe have shown that the expansion is accelerating; indeed, these astronomical observations were the source of the 2011 Nobel Prize in Physics. We don't know what is causing this, and have given the name "dark energy" to whatever it is that is causing it. The simplest explanation for dark energy is that it is vacuum energy. Measurements from cosmology indicate a vacuum energy density corresponding to about 10^{-29} grams per cubic centimeter. That is, the energy density of vacuum energy is 29 orders of magnitude less than the mass-energy density of water. Obviously, we can ignore this in our every day life. However, if you look at the Universe as a whole, most of it is empty; our planet is a very special place that is, compared to most of the Universe, extremely dense with regular atoms. In the Universe as a whole, dark energy makes up three quarters of the energy density. Even though this density may be 120 orders of magnitude smaller than what naive estimates from our theory would suggest, it is coming to dominate the evolution of our Universe.

This page titled 11.4.2: Vacuum Energy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

12: Multiple Particle States

12.1: Indistinguishable Particles
12.2: Notating Multiple Particle States
12.3: The Exchange Operator
12.4: Fermions and Bosons
12.5: The Pauli Exclusion Principle
12.6: Entangled Particles

This page titled 12: Multiple Particle States is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

12.1: Indistinguishable Particles

Every electron is exactly the same as every other electron. Thus, all electrons are indistinguishable. This means that if you have a state with two electrons, you can swap the two electrons and it cannot change anything physically observable from that state.

To make this concrete, suppose the state $|\psi\rangle$ is a state with two electrons. Let's define $|\psi'\rangle$ as the state with those two electrons swapped. Then, the expectation value of any operator must be the same for these two different states:

$$\langle\psi|\hat{\mathcal{O}}|\psi\rangle = \left\langle\psi'|\hat{\mathcal{O}}|\psi'\right\rangle \tag{12.1}$$

Also, the probability for any measurement of any observable to be made must be the same for the two states. That is, if $\langle\phi|$ is an eigenstate of a given observable, then

$$|\langle\phi \mid \psi\rangle|^2 = |\langle\phi \mid \psi'\rangle|^2 \tag{12.2}$$

If you think about it, however, this does not mean that the two states must be identical! However, they must be close enough such that anything physically observable from the state must be identical. Below, we will introduce the exchange operator as a way of quantifying the effect of identical particles on quantum states.

This page titled 12.1: Indistinguishable Particles is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

12.2: Notating Multiple Particle States

Before we go further, we need to refine our notation so that we can keep track of two different particles. We can construct a two-particle state by putting together two states for each individual particle with:

$$|\psi_1\rangle \otimes |\phi_2\rangle \tag{12.3}$$

The ⊗ operator indicates that we're putting these two states together to form a composite state. It's sometimes called a "direct product", but it's not really all that much like multiplication. Really, it just means that we're making some composed state that combines particle 1 in state $|\psi\rangle$ and particle 2 in state $|\phi\rangle$. The subscript indicates which particle we're talking about; the rest of the stuff inside the ket indicates the state of that particular particle.

For simplicity, we will often omit the ⊗ symbol in the "direct product", and just write the two states next to each other, e.g.

$$|\psi_1\rangle\,|\phi_2\rangle \tag{12.4}$$

Again, this does not mean that we're multiplying two ket vectors, which is something we can't do. Instead, it means that we're composing the states. If these were spin states, we would not represent this with two column vectors. Instead, we'd represent it with a single four-row column vector; the first two rows have the column vector representation of whatever state the first particle is in, and the second two rows have the column vector representation of whatever state the second particle is in.

If an operator operates on this state, it only affects the state for the particle it is an operator for. That is, if "spin-z for particle 2" is the observable we're talking about, then the operator $\hat{S}_{z2}$ only operates on (in this example) the state $|\phi_2\rangle$. Indeed, you can treat $|\phi_1\rangle$ as if it were a constant:

$$\hat{S}_{z2}\,|\psi_1\rangle\,|\phi_2\rangle = |\psi_1\rangle\,\hat{S}_{z2}\,|\phi_2\rangle \tag{12.5}$$

As an example, suppose that particle 1 is in the state $|+z\rangle$ and particle 2 is in the state $|-z\rangle$. If we apply the $\hat{S}_{z2}$ operator to this state, we get:

$$\begin{aligned}\hat{S}_{z2}\,|+z_1\rangle\,|-z_2\rangle &= |+z_1\rangle\,\hat{S}_{z2}\,|-z_2\rangle \\ &= |+z_1\rangle\left(\frac{-\hbar}{2}\right)|-z_2\rangle \\ &= \left(-\frac{\hbar}{2}\right)|+z_1\rangle\,|-z_2\rangle\end{aligned} \tag{12.6}$$

Here, we have taken advantage of the fact that $|-z_2\rangle$ is an eigenstate of $\hat{S}_{z2}$, and replaced the action of the operator with a simple multiplication by the eigenvalue.

There will be some operators (e.g. the forthcoming exchange operator) that don't operate on just one of the two particles, but on both at the same time.

Similarly, with inner products, bra versions of a state only "stick" to ket versions of a state on the straight side of the bra-ket notation if they are states for the same particle. Thus, suppose we had a composite state:

$$|\xi\rangle = |\psi_1\rangle\,|\phi_2\rangle \tag{12.7}$$

The corresponding bra vector is:

$$\langle\xi| = \langle\psi_1|\,\langle\phi_2| \tag{12.8}$$

Normalization of this state is then expressed as:

$$\begin{aligned}\langle\xi\,|\,\xi\rangle &= \left(\langle\psi_1|\,\langle\phi_2|\right)\left(|\psi_1\rangle\,|\phi_2\rangle\right) \\ &= \langle\psi_1\,|\,\psi_1\rangle\,\langle\phi_2\,|\,\phi_2\rangle \\ &= 1\end{aligned} \tag{12.9}$$

We've rearranged states here a bit. We moved the $|\psi_1\rangle$ from after the $\langle\phi_2|$ to before it. This should make you a little nervous; we've seen that with matrices and other things that aren't simple numbers, multiplication is not necessarily commutative. However, again, in this case, when it comes to inner products, a state for a different particle can be treated as a constant with respect to inner

products for the first particle. As such, it's entirely legitimate to move $|\psi_1\rangle$ into, out of, and through inner products on particle 2 (at least in the case of the simple composed states we're talking about here).

This page titled 12.2: Notating Multiple Particle States is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

12.3: The Exchange Operator

The exchange operator, notated here by $\hat{P}_{12}$, just exchanges particle 1 for particle 2. In order to satisfy the conditions described in Section 12.1, a state composed of two indistinguishable particles (e.g. two electrons) must be an eigenstate of the exchange operator. Suppose that $|\xi\rangle$ is such a state. This means that

$$\hat{P}_{12}|\xi\rangle = c|\xi\rangle \tag{12.10}$$

where c is the eigenvalue. Suppose that we apply the exchange operator twice. What will happen? We should get back to the original state! We've just swapped the two particles back. Let's apply this twice:

$$\begin{aligned} \hat{P}_{12}\hat{P}_{12}|\xi\rangle &= \hat{P}_{12}(c|\xi\rangle) \\ &= c\hat{P}_{12}|\xi\rangle \\ &= c^2|\xi\rangle \end{aligned} \tag{12.11}$$

If the result of applying this exchange operator twice must be the state we started with, then we must have $c^2 = 1$. This is regular old fashioned squaring, not taking the absolute square. That $c^2 = 1$ means that there are only two possibilities for the eigenvalue of the exchange operator working on a state of two indistinguishable particles: $c = 1$ or $c = -1$.

This page titled 12.3: The Exchange Operator is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

12.4: Fermions and Bosons

In quantum mechanics, there are two kinds of particles. **Fermions** are particles that are antisymmetric under the exchange operator; that is, if $|\xi\rangle$ is a two-particle state for two indistinguishable fermions, $\hat{P}_{12}|\xi\rangle = -|\xi\rangle$. **Bosons** are particles that are symmetric under the exchange operator; that is, if $|\xi\rangle$ is a two-particle state for two indistinguishable bosons, $\hat{P}_{12}|\xi\rangle = |\xi\rangle$. This is summarized below:

$$\hat{P}_{12}|\xi\rangle = \begin{cases} |\xi\rangle \text{ for a two-boson state} \\ -|\xi\rangle \text{ for a two-fermion state} \end{cases} \tag{12.12}$$

Which particles are which? Particles that have half-integral spin— which includes the spin-1/2 electrons we've been talking about all this time— are fermions. Other fermions include protons, neutrons, quarks, and neutrinos. Particles with integral spin are bosons. Bosons include photons, pions, and the force carriers for the weak and strong nuclear forces.

How do you create a two-fermion state with a total z component of angular momentum equal to zero? The most obvious first thing to guess is just to assign each particle angular momentum in a different direction, so that they cancel:

$$|\xi\rangle = |+z_1\rangle\,|-z_2\rangle \tag{12.13}$$

However, this state doesn't work! Why not? Consider the operation of the exchange operator on it:

$$\hat{P}_{12}\,|+z_1\rangle\,|-z_2\rangle = |+z_2\rangle\,|-z_1\rangle \tag{12.14}$$

We started with particle one having positive z-spin and particle 2 having negative z-spin. After the exchange, it's the other way around. However, this isn't the same state, nor is it a constant times the original state. On other words, this state is not an eigenstate of the exchange operator. Therefore, it's not a valid quantum state if particle 1 and particle 2 are indistinguishable particles (e.g. if they're two electrons).

A valid two-fermion spin state with total angular momentum zero would be:

$$|\xi\rangle = \frac{1}{\sqrt{2}}|+z_1\rangle\,|-z_2\rangle - \frac{1}{\sqrt{2}}|+z_2\rangle\,|-z_1\rangle \tag{12.15}$$

To verify that this works, let's try the exchange operator on this state:

$$\begin{aligned} \hat{P}_{12}|\xi\rangle &= \hat{P}_{12}\left(\frac{1}{\sqrt{2}}|+z_1\rangle\,|-z_2\rangle - \frac{1}{\sqrt{2}}|+z_2\rangle\,|-z_1\rangle\right) \\ &= \frac{1}{\sqrt{2}}\hat{P}_{12}\,|+z_1\rangle\,|-z_2\rangle - \frac{2}{\sqrt{2}}\hat{P}_{12}\,|+z_2\rangle\,|-z_1\rangle \\ &= \frac{1}{\sqrt{2}}|+z_2\rangle\,|-z_1\rangle - \frac{1}{\sqrt{2}}|+z_1\rangle\,|-z_2\rangle \\ &= -|\xi\rangle \end{aligned} \tag{12.16}$$

Sure enough, this state is an eigenstate of the exchange operator. What's more, the eigenvalue is −1, which is required for fermions. (If you're wondering about why we mess about with all of the $1/\sqrt{2}$ coefficients, we do that so that $|\xi\rangle$ is properly normalized. You can verify that this is the case, and indeed doing so would be good practice in doing algebra with bra and ket vector representations of multiple particle states.)

This page titled 12.4: Fermions and Bosons is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

12.5: The Pauli Exclusion Principle

The Pauli Exclusion Principle states that no two fermions may occupy the same quantum state. This principle is absolutely crucial to life as we know it; without it, we would not have the Periodic Table of chemistry, nor would we have a lot of the rest of the structure of matter. This doesn't mean, however, that only one electron in the Universe is allowed to have positive z spin! Obviously, we have many more than two electrons in the Universe. However, if you have a quantum state, such as an energy level in an atom, where you can put electrons, you can only put two electrons into that energy level. Why two, and not one? Because of electron spin; as long as the two electrons have opposite spin (or, more precisely, are in a combined spin state with spin angular momentum zero such that they are antisymmetric under exchange), then you can put two electrons into the same state. It is possible to have two electrons with the same spin, so long as something else is different about their quantum states. So, for example, you could have two electrons with the same spin if they were in different orbitals in an atom.

Why can't you put more than one fermion in the same state? Because it's impossible to construct an antisymmetric state vector two fermions in the same state. Suppose you have a state $|\psi\rangle$, and you want to put two fermions into it. We know that the state:

$$|\psi_1\rangle\,|\psi_2\rangle \tag{12.17}$$

won't work, because the exchange operator working on it just produces the same state back, not the negative of the same state:

$$\hat{P}_{12}\,|\psi_1\rangle\,|\psi_2\rangle = |\psi_2\rangle\,|\psi_1\rangle = |\psi_1\rangle\,|\psi_2\rangle \tag{12.18}$$

This is an eigenvalue of the exchange operator, which is good, but the eigenvalue is +1. This would work for bosons; indeed, because of this, you can put as many bosons as you want all into the same state. However, for fermions, the eigenvalue of the exchange operator working on the two-particle state needs to be −1. If we try to construct an antisymmetric wave vector with both of these electrons in the same state:

$$\frac{1}{\sqrt{2}}|\psi_1\rangle\,|\psi_2\rangle - \frac{1}{\sqrt{2}}|\psi_2\rangle\,|\psi_1\rangle \tag{12.19}$$

we just end up with 0, which isn't a state at all. Thus, if you have two indistinguishable fermions, there must be something different about their states; you can't put more than one fermion into a single quantum state.

This page titled 12.5: The Pauli Exclusion Principle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

12.6: Entangled Particles

When two particles' quantum state is a combined quantum state, we say that those two particles are entangled. Most of the time we encounter such states, we don't worry about it too much. The two electrons in the ground state of Helium have entangled states, because they are indistinguishable particles. You can't talk about the state of one electron without talking about the state of another.

Entangled quantum states become more interesting when you separate the two particles. Suppose that there is some sort of reaction that produces two electrons that have a total spin angular momentum of zero. We've seen before that the state of these two electrons is then:

$$\frac{1}{\sqrt{2}}|+z_1\rangle\,|-z_2\rangle - \frac{1}{\sqrt{2}}|+z_2\rangle\,|-z_1\rangle \tag{12.20}$$

Although the total z angular momentum of this combined state is 0, a definite value, the angular momentum of an individual electron is not in a definite state. Now suppose that you separate these two electrons; it may be that the reaction that produces them sends them shooting off in two directions, which for discussion purposes we shall call "left" and "right".

Now let's suppose that somebody far off to the left detects the left electron and measures its z-spin. This measurement will collapse the wave function of the left electron, putting it into a state of definite z spin. However, because it's a combined state for the two electrons, you can't collapse the wave function of just one of them; you have to collapse the entire state all at once. Therefore, if somebody measures the z spin of the left electron, the wave function of the right electron also collapses at that moment, even if nobody has made a measurement on it. If the left observer measures that the left electron is spin up, then anybody off to the right will observe that the right electron is spin down; the right electron is no longer in an indefinite state, even though nothing was done to it.

This behavior of entangled particles is what Einstein referred to as "spooky action at a distance". (citation needed.) Not only was he disturbed by the stochastic nature √ of quantum mechanics, he was also bothered by what seemed to be communication faster than the speed of light. Does some sort of signal traverse from one electron to the other electron in order to communicate the fact that their mutual wave function has collapsed? Together with two other physicists, Podolsky and Rosen, Einstein argued that this behavior indicated that quantum theory had to be incomplete. In 1935, they published a paper describing what is now known as the "EPR Paradox" (Einstein et al., 1935). If quantum mechanics is indeed incomplete, then there would need to be some sort of "local hidden variable" that tells a particle which way its wave function should collapse when that particle is measured. This variable is "hidden" because it is not accounted for in quantum mechanics. In the early 1960's, physicist John Bell proposed experiments that would test the EPR paradox by being able to tell the difference between the standard predictions of quantum mechanics and the predictions of a theory that had some sort of local hidden variables (citation needed). √ Experiments performed since then have shown that in fact standard quantum mechanics does predict the correct results, and that therefore there are no local hidden variables. The fact is that, somehow, the wave function of an electron can collapse when another electron is measured— and that other electron may, at least in principle even if this is not realizable in practice, be light-years away. This raises philosophical issues associated with the interpretation of quantum mechanics, but also indicates that quantum mechanics remains a very robust theory.

This page titled 12.6: Entangled Particles is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

13: The Schrödinger Equation

13.1: Where we are so far
13.2: Stating the Equation
13.3: Free Particles and the de Broglie Wavelength
13.4: Quantized Energy Levels in Bound Systems
13.5: The Simple Harmonic Oscillator
13.6: The Hydrogen Atom
13.7: Interpretation of the Wave Function ψ(x)

This page titled 13: The Schrödinger Equation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

13.1: Where we are so far

We have focused primarily on electron spin so far because it's a simple quantum system (there are only two basis states!), and yet it still shows much of the peculiar nature of reality on the quantum level. In particular, we've seen the following things in the theory of quantum mechanics:

- A "system" (e.g. the angular momentum vector of an electron) may be an an indefinite state, also sometimes called a "mixture" of states, where an observable doesn't have a set value. Rather, the state of the system is such that if the observable were measured, there is a probability of different values being observed. The mathematical theory represents this by allowing states to be sums of coefficients times orthogonal basis states. For example, with angular momentum of a spin-1/2 particle such as an electron, the basis states are $|+z\rangle$ and $|-z\rangle$.
- Observables may take on quantized values. For example, every time you measure the z component of angular momentum of an electron, you get either $+\hbar/2$ or $-\hbar/2$. This is in sharp contrast to what you'd see in classical physics.
- What propagates in quantum mechanics is amplitudes. For example, if an electron is in state $|\psi\rangle$, the amplitude to measure it to have z angular momentum $+\hbar/2$ is $\langle +z \,|\, \psi\rangle$. The probability, which is what we can really find in experiments, is the absolute square of the amplitude; in this example, that would be $|\langle +z \,|\, \psi\rangle|^2$.
- Different observables may be orthogonal (the second use of this term). If they are, then a system can not be in a definite state for those two observables at the same time. The projections of angular momentum along different axes are orthogonal; position and momentum along the same direction are orthogonal.
- Observables in quantum mechanics are paired with operators. A quantum mechanical operator operates on a quantum state (represented by a ket vector), and the result of that operation is another (non-normalized) quantum states (i.e. another ket vector). For example, if we call the z component of angular momentum spin-z or just s_z, the operator that goes with it is $\hat{S}_z$, the spin-z operator. Operators are quite abstract, and form a mathematical part of the theory that is useful, but is difficult to interpret and associate directly with something that you could observe.
- A state that is a definite state for a given observable is an eigenstate of that operator. (We would also say that the ket vector that represents that state is an eigenvector of the operator; if we're representing operators as matrices, then the column vector that represents the state is an eigenvector of the operator.) An operator working on one of its eigenstate returns a constant times the same state. That constant is called the eigenvalue associated with the eigenstate. If this operator corresponds to an observable, that eigenvalue must be a real number, and corresponds to the physical measurement you'd make of that observable. For example:

  $$\hat{S}_z|+z\rangle = \frac{\hbar}{2}|+z\rangle \tag{13.1}$$

  This equation is the eigenvalue equation, in this case specifically for the z-spin operator and the $|+z\rangle$ state. The state $|+z\rangle$ is a state of definite z-spin, so it is an eigenstate of the z-spin operator $\hat{S}_z$. The eigenvalue equation for this state and this operator includes the constant $\hbar/2$, which is the actual value of the z component of spin angular momentum that an electron in state $|+z\rangle$ has.
- You can find the expectation value for a system in state $|\psi\rangle$ for a given observable by sandwiching the observable's operator between $\langle\psi|$ and $|\psi\rangle$. For example, the expectation value for z-spin for a given electron is $\left\langle \psi \middle| \hat{S}_z \middle| \psi \right\rangle$. The expectation value is the average value you'd get if you measured the observable for that state. That is, if you took a large number of systems in that state and measured the observable for all of those systems, you'd get different results, with probabilities for each result predicted by the mathematics of quantum mechanics. The average of all those results would be the expectation values.

Although the eigenvalue equation is fairly abstract, it's a very important part of the mathematical theory of quantum mechanics. The only direct connection it has to what we might observe in the lab is that it extracts (in the form of the eigenvalue) the measured quantity for the observable that you'd get for a given eigenstate (i.e. definite state) of that observable's operator. However, the operator itself doesn't represent any particular physical operation you might perform in the lab.

From a broader point of view, the eigenvalue equation is the equation you can use to figure out what states are possible definite states for a given operator, and therefore what values you might measure for the observable associated with that operator.

This page titled 13.1: Where we are so far is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

13.2: Stating the Equation

As was briefly mentioned in Section 8.5, the energy operator is an operator so important to quantum mechanics that it gets its own proper name, the Hamiltonian, and the eigenvalue equation for it also gets its own name, the Schrödinger equation. It is this equation that allows us to figure out the energy states of a system, and it could be argued that energy states are the most important states in quantum mechanics. It is energy levels in atoms that provides all of the structure that gives us the Periodic Table of the Elements, and it is transitions between those energy levels that we observe in a number of both terrestrial and astrophysical contexts.

At the most base level, the Schrödinger equation is just the energy operator eigenvalue equation:

$$\hat{H}|\psi\rangle = E|\psi\rangle \tag{13.2}$$

$\hat{H}$ is the Hamiltonian; it's the operator that corresponds to energy as an observable. Solutions $|\psi\rangle$ to this equation are the eigenstates of energy. The value E that goes with a given solution $|\psi\rangle$ is the energy associated with that state. Technically, this equation is called the Time-Independent Schrödinger Equation. (There is also a full Schrödinger equation that describes how quantum states evolve in time.)

A full investigation of the Hamiltonian requires differential calculus, so we won't fully present it here. However, you can break the Hamiltonian into two parts. In doing so, we're going to go to a wave function representation of the state vector $|\psi\rangle$. Whereas we have used column vectors to represent spin states, it is more traditional (and more useful) to represent energy states as functions of position $\psi(x, y, z)$. As with a regular function, $\psi(x, y, z)$ is just something into which you can plug a position (i.e. values of x, y, and z) and get a number— although in this case that number can be a complex number. Dividing the Hamiltonian into two parts and writing the state as a wave function yields this form of the Schrödinger equation:

$$\hat{K}\psi(x, y, z) + V(x, y, z)\psi(x, y, z) = E\psi(x, y, z) \tag{13.3}$$

The Hamiltonian here has been divided into the kinetic energy operator $\hat{K}$, and the potential energy $V(x, y, z)$. Note that the action of the potential energy operator is just multiplying the wave function by the potential energy! The aforementioned differential calculus is buried inside $\hat{K}$. In fact, there isn't a single Schrödinger equation. Rather, there's a different one for each form of the potential $V(x, y, z)$. This also means that the solutions $\psi(x, y, z)$ will be different for each potential.

Although the full time-independent Schr¨odinger equation is in fact a function in the full 3-D space that we live in, for much of what we do below we will simplify it and consider only one dimensional systems. This makes dealing with it conceptually simple, but does not obscure any of the essential physical results. Such systems can in fact be realistic. For example, if you consider a mass moving on a spring attached to a wall, that is essentially a one-dimensional system, as the mass moves only forwards and backwards along the direction the spring is oriented.

This page titled 13.2: Stating the Equation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

13.3: Free Particles and the de Broglie Wavelength

A particle is called a "free particle" if its potential is constant. That is, there are no potential energy wells or barriers anywhere. It's simplest to choose that constant potential energy to be zero, as that reduces the Schrödinger equation to:

$$\hat{K}\psi(x) = E\psi(x) \tag{13.4}$$

(in the one-dimensional case). Solutions to this equation are called "plane-wave" solutions. They are states with definite momentum $p = E^2/2m$ (which is exactly what you'd expect if you compare momentum and kinetic energy in classical physics). However, their position is completely undetermined; there is equal probability for any x, which is what you'd expect for a state of definite momentum given the Heisenberg Uncertainty Principle. The functional form of $\psi(x)$ is just a standard wave:

$$\psi(x) = A\cos(2\pi x/\lambda) \pm iA\sin(2\pi x/\lambda) \tag{13.5}$$

where A is a constant (a complex number) that normalizes the wave. The ± depends on whether the wave is moving to the right (i.e. momentum is in the $+x$ direction) or to the left (i.e. momentum is in the $-x$ direction). The normalization condition will only put a constraint on the absolute square of A, meaning that there will be many complex numbers that satisfy it. As such, there isn't one single solution to this equation; however, all of the solutions do give the same predictions for measurable things such as the probability of finding the electron at a given spot. The value λ that shows up in these equations is the wavelength; that is it's the range of x over which it takes the sine or the cosine to go through one complete cycle. Note that although $\psi(x)$ varies with space, the probability of finding x at any given position, $|\langle x \mid \psi\rangle|^2$, does not! See Section 13.7 for more details about this.

In this case, the energy levels are not quantized. E can be anything in the equation above. A different energy E does correspond to a different wavelength in the plane wave represented by ψ. In these solutions, the energy E of the particle is related to the wavelength λ of the wave function by:

$$E = \frac{h^2}{2m\lambda^2} \tag{13.6}$$

It's more traditional to express this wavelength, called the de Broglie wavelength, in terms of the momentum of the particle:

$$\lambda = \frac{h}{p} \tag{13.7}$$

You can get this equation directly from the previous equation by using the relationship $E = p^2/2m$, that results from the combination of kinetic energy $E = \frac{1}{2}mv^2$ and momentum $p = mv$. The constant h here is a version of Plank's Constant, related to $\hbar$ by $h = 2\pi\hbar$.

For example, what is the de Broglie wavelength of an electron moving at 1×10^6 m/s (a "typical" speed for an atomic electron)? We would plug the right numbers into this equation:

$$\begin{aligned}\lambda &= \frac{h}{p} = \frac{h}{mv} \\ &= \frac{6.626\times10^{-34}\ \text{kg m}^2\ \text{s}^{-1}}{\left(9.109\times10^{-31}\ \text{kg}\right)\left(1\times10^{6}\ \text{m s}^{-1}\right)} \\ &= 7\times10^{-10}\ \text{m} = 0.7\ \text{nm}\end{aligned} \tag{13.8}$$

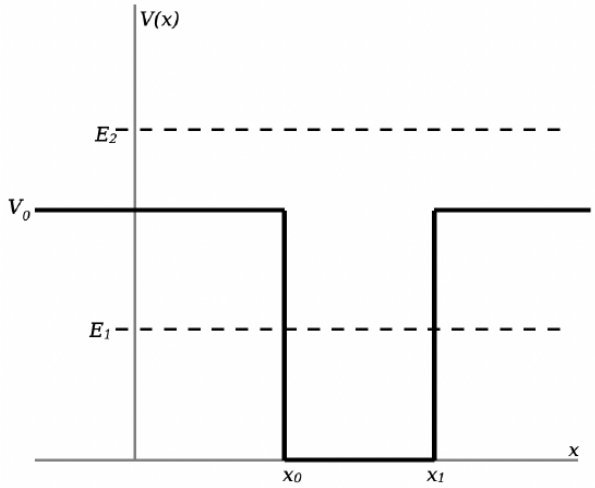
For comparison, this is about 1/1000 the wavelength of visible red light.

Many of the physical effects peculiar to quantum mechanics show up as wave interference between different components of a wave function $\psi(x)$. All waves, including those that derive from classical physics (such as waves on a string, sound waves, or electromagnetic (i.e. light) waves), show interference effects. The fact that the wave function, this abstract mathematical object which is used to figure out things about the state of a particle, also shows interference effects is what we mean when we say that sometimes particles behave like waves. In general, the longer the wavelength of a wave (i.e. the larger λ is), the easier it is to see interference effects. The de Broglie wavelength indicates that wavelength is inversely proportional to momentum. For a non-relativistic particle (which is implied here, as the Schr¨odinger equation assumes non-relativistic particles), $p = mv$. Thus, for particles moving at a given velocity, the larger m is, the smaller λ is. This is why it is so difficult to observe quantum interference effects for larger objects; the effective wavelength, and thus the typical separations that you'd need to see those effects, becomes tiny.

This page titled 13.3: Free Particles and the de Broglie Wavelength is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

13.4: Quantized Energy Levels in Bound Systems

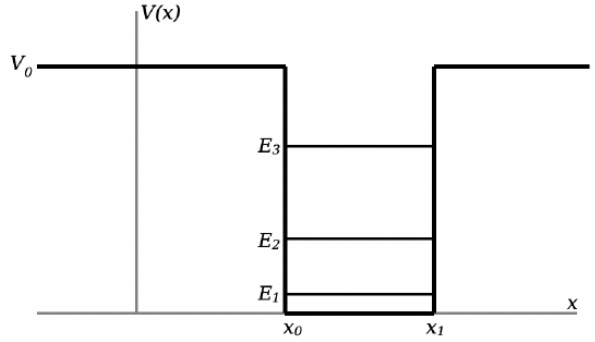
Whereas a free particle has a continuum of energy states available to it, if the particle is bound in a potential its available energy states are quantized. For a particle to be bound in a potential, the potential must limit a region of space that the particle could access classically. That is, if there is an energy well of some depth, the particle's energy over the bottom of the well must be less than the depth of the well:



A finite square potential well, with potential energy 0 chosen to be the bottom of the well. (Sometimes, one chooses the top of the well to be potential energy 0.) If a particle's total energy is less than the depth of the well V_0, as is the case for energy E_1 in the figure, that particle is bound in the well. If the particle's total energy is greater than the depth of the well, as is the case for energy E_2 in the figure, that particle is free, although its wave function is still influenced by the presence of the well.

In this example, a particle would be classically limited to the range of position $x_0 < x < x_1$ (see Section 2.2.2). Quantum mechanically, the particle is most likely to be found in that range, but it turns out there is a finite non-zero probability that the particle is found outside the classically allowed volume!

In this square well, there are a finite number of energy states available to a particle. Classically, a bound particle could have any energy $0 \leq E < V_0$. However, quantum mechanically, the particle must be in one of specific separated states. This is analogous to what we've seen with angular momentum of a spin-1/2 particle, where the projection of angular momentum along one axis must take on one of two specific separated values. For the square well, the allowed energy levels qualitatively looks something like:



In this example, there are only three allowed energy states, and the energies of those states are E_1, E_2, and E_3. Notice that the lowest allowed energy level is not 0! That is, you can't have a particle that's got no kinetic energy. This fits with the Heisenberg Uncertainty Principle. By binding the particle into the potential well, you've set constraints on the particle's position: it's most likely to be between x_0 and x_1. The particle isn't equally probable to be anywhere, so the uncertainty on the particle's position Δx is finite (and indeed will be something close to $x_1 - x_0$). As such, there must be a corresponding uncertainty in the particle's momentum Δp, so it's impossible for the particle to be in a zero-momentum state.

This page titled 13.4: Quantized Energy Levels in Bound Systems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available

upon request.

https://phys.libretexts.org/@go/page/56888

13.5: The Simple Harmonic Oscillator

One important potential energy function is the Simple Harmonic Oscillator, or SHO. This is the potential energy of a spring (so long as you don't stretch of squish the spring too much). It also turns out to be a decent approximation, at least for lower energy levels, for a number of quantum systems. One such system is the vibrational energy states of a Hydrogen molecule H_2. The form of this potential, in one dimension, is:

$$V(x) = \frac{1}{2} m\omega^2 x^2 \tag{13.9}$$

Here, m is the mass of the particle moving in the potential. ω is the "natural frequency of oscillation" for the potential; for a classical spring, it would correspond to $2\pi/T$, where T is the period of oscillations. (Of course, for a classical spring, the system could also have any energy!)

The solution to the one dimensional Schrödinger equation for this potential gives the following energies for the energy eigenstates:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \tag{13.10}$$

where n is an integer 0, 1, 2, ... As written, this potential is an infinitely high potential $(V(x)$ just keeps going up as x gets farther and farther from 0.) As such, there are an infinite number of allowed energy levels. Of course, as an approximation to a real physical system, usually the approximation will get worse and worse as x gets farther and farther from 0, which means that the solutions less and less of a good approximation to the real energy system for higher and higher energy levels.

This page titled 13.5: The Simple Harmonic Oscillator is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

13.6: The Hydrogen Atom

The potential that an electron in a Hydrogen atom experiences results from the electrostatic interaction between the electron (which is negatively charged) and the proton (which is positively charged). (A Hydrogen nucleus is composed of a single proton.) It is traditional (and convenient) to choose the zero level of the potential to be when the electron is extremely far away from the proton. This means that the potential energy gets more and more negative as the electron gets closer and closer to the proton. The form of this potential is:

$$V(r) = -\frac{1}{4\pi\epsilon_0}\frac{e^2}{r} \tag{13.11}$$

In this equation, ϵ_0 is a fundamental constant related to the effective strength of the electromagnetic force; it's value is 8.854 × 10^{-12} C^2 m^{-1}J. The letter e indicates the elementary charge; it is the charge on the proton, and the absolute value of the charge on the electron. It's value is 1.602 × 10^{-19}C, where C is "Coulombs", the SI unit of charge. Finally, r indicates the distance between the electron and the proton. If we consider the proton to be at the origin, then $r = \sqrt{x^2+y^2+z^2}$, with (x, y, z) indicating the position of the electron. The Schrödinger equation for an electron in a Hydrogen atom is then:

$$\hat{K}\psi(\vec{r}) - \frac{1}{4\pi\epsilon_0}\frac{e^2}{r}\psi(\vec{r}) = E\psi(\vec{r}) \tag{13.12}$$

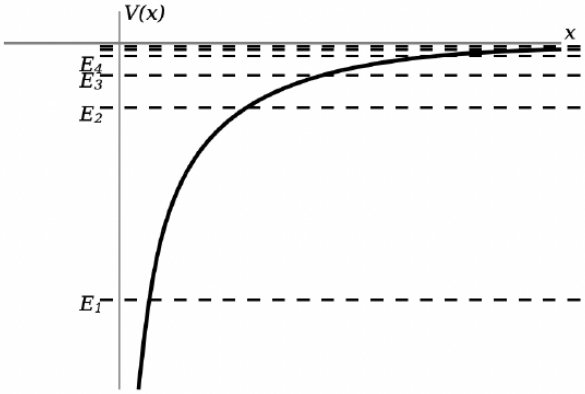
Here, we've written $\psi(\vec{r})$ as a shorthand for $\psi(x, y, z)$; the vector form of $r, \vec{r}$, indicates that the wave function depends on the electron's displacement from the origin. In fact, it's more common and more practical to express the wave function in terms of spherical coordinates (r \theta, \phi), where r is the distance from the origin, θ is the angle off of the z axis, and ϕ is the angle off of the x axis in a projection on to the $x - y$ plane. (Even though the potential is spherically symmetric, in that it only depends on the distance from the origin, it turns out that the wave function solution is not always spherically symmetric. The reason for this is angular momentum, and that will be discussed in the next section.)

As with the other potentials discussed above, the allowed energy levels are quantized. In the Hydrogen atom, that energy level is specified by a principle quantum number n. The energy levels in a Hydrogen atom are:

$$E_n = \frac{-13.6\text{eV}}{n^2} \tag{13.13}$$

The ground state of Hydrogen has an energy of −13.6 eV. That is, if you want to tear the electron off of the Hydrogen atom (a process known as ionization), you need to somehow provide at least 13.6 eV of energy in order to give the electron enough energy to make it out to extremely large distances away from the proton.

The figure below slows the Hydrogen atom potential and the first few energy levels. Higher and higher energy levels are less tightly bound (it takes less energy to ionize the atom, freeing the electron). They also get closer and closer together.



In fact, there are three quantum numbers associated with the solution to the Hydrogen atom. In addition to the principle quantum number n, there is also the total orbital angular momentum quantum number l, and the orbital z-angular momentum quantum number m. You could notate the energy eigenstates in a Hydrogen atom, corresponding to a state that the electron could actually be in, by $|n, l, m\rangle$. The ground state for a Hydrogen atom, in this notation, would be $|1, 0, 0\rangle$. For actual electrons, there's a fourth quantum number you have to specify: s, the z-spin of the electron. For every state n, l, m in a Hydrogen atom, there are in fact two

different electron states, one where the electron has z-spin $+\hbar/2$, one where the electron has z-spin $-\hbar/2$. We'll call this quantum number s, and it will have either the value $+1/2$ or $-1/2$.

In Hydrogen, to very good approximation the energy of an eigenstate $|n,l,m\rangle$ only depends on the principal quantum number n, regardless of l and m. That is, the eigenvalue associated with states $|2,1,0\rangle$ and $|2,0,0\rangle$ are exactly the same (and are equal to $E=\big(-13.6\text{eV}\big)/n^2=-3.4\text{eV}\big)$. In atoms with more than one electron, things get more complicated, and the energy of a given state may depend on the other quantum numbers.

This page titled 13.6: The Hydrogen Atom is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

13.7: Interpretation of the Wave Function ψ(x)

In general, it is best to view $\psi(x)$ the same way that you view $|\psi\rangle$. It's an abstract mathematical object that represents the state of the system. Quantum mechanics is then a theory, a mathematical model of reality that includes rules for manipulating $\psi(x)$ (or other representations of $|\psi\rangle$) in order to make predictions about the results of experiments, such as probabilities for observing particles in certain states, or expectation values for certain values.

It turns out that there is one particularly simple rule that can be applied to $\psi(x)$ in order to learn something about the state of the system. if $\psi(x)$ is a properly normalized single-particle wave function, then the construction $\psi^*(x)\psi(x)dx$ is the probability of finding that particle between position x and position $x+dx$, where dx is a small range of x. (By "small", we mean small enough that $\psi(x)$ does not appreciably change over the range.) As an example, consider the free particle wave function:

$$\psi(x) = A[\cos(2\pi x/\lambda) \pm i\sin(2\pi x/\lambda)] \tag{13.14}$$

If we want to find the probability for finding a particle at a given position, we multiply this function by its complex conjugate:

$$\begin{aligned}
\psi^*(x)\psi(x)dx &= A^*A\left[\cos\left(\frac{2\pi x}{\lambda}\right) \mp i\sin\left(\frac{2\pi x}{\lambda}\right)\right)\left(\cos\left(\frac{2\pi x}{\lambda}\right) \pm i\sin\left(\frac{2\pi x}{\lambda}\right)\right]dx \\
&= A^*A\left[\cos^2\left(\frac{2\pi x}{\lambda}\right) + \sin^2\left(\frac{2\pi x}{\lambda}\right)\right. \\
&\qquad \left. \pm i\cos\left(\frac{2\pi x}{\lambda}\right)\sin\left(\frac{2\pi x}{\lambda}\right) \mp i\cos\left(\frac{2\pi x}{\lambda}\right)\sin\left(\frac{2\pi x}{\lambda}\right)\right]dx \\
&= A^*A\left[\cos^2\left(\frac{2\pi x}{\lambda}\right) + \sin^2\left(\frac{2\pi x}{\lambda}\right)\right]dx
\end{aligned} \tag{13.15}$$

To simplify this further, we can use the trigonometric identity $\sin^2\phi + \cos^2\phi = 1$ (this applies for all ϕ). Thus, we are left with:

$$\psi^*(x)\psi(x)dx = |A|^2dx \tag{13.16}$$

That is, the probability of finding a free particle at any x within a given range dx is always the same. This correspoinds to an infinite uncertainty in position x, which is what we need given that this state has a definite momentum $p = h/\lambda$.

The construction $\psi^*(x)\psi(x)dx$ works for any one-dimensional wave function for calculating the probability of finding the particle at a given position. Using the three dimensional version of this construction on the solutions to the Hydrogen atom is what gives us the "electron cloud" diagrams you may have seen for electron orbitals. More about that in the next chapter.

This page titled 13.7: Interpretation of the Wave Function ψ(x) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

14: Atomic Orbitals

In this chapter, we will be talking primarily about the Hydrogen atom. However, a lot of what we are saying will be an approximation to other atoms. In the Hydrogen atom, there are just two particles interacting, an electron and a proton. The proton is more than a thousand times more massive than the electron. That means that as a result of their interaction, the electron moves far more than the proton does. As such, we can view it as an electron moving about in the potential of the proton. Except for the fact that it provides this potential, to first order we can ignore the proton, and just consider the electron as the quantum mechanical particle.

For other atoms, there are additional electrons. The energy levels that we have calculated for Hydrogen can very easily be adapted to take into account a nucleus with more protons— all you have to do is multiply all of the energy levels by Z^2, the square of the number of protons in the nucleus (which is also the positive charge of the nucleus). We can then approximate other atoms by putting electrons into all of these energy levels. Because electrons are fermions, we can only put two into any given orbital. (Two because there are two possible spin states for an electron.) However, this implicitly assumes that the electrons are interacting only with the nucleus, and not with each other. That approximation will allow us to get a lot of insight into the structure of (for example) the Periodic Table of the elements, but is too much of an approximation to be able to figure out precise energy levels.

14.1: The Schrödinger Equation

14.2: The Orbitals

14.3: Visualizing Orbitals

14.3.1: s Orbitals

14.3.2: p Orbitals

14.3.3: d Orbitals

This page titled 14: Atomic Orbitals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

14.1: The Schrödinger Equation

The one-dimensional Coulomb potential was presented in section 13.6. Of course, the real Hydrogen atom is three-dimensional. The Schr¨odinger equation that results is almost identical:

$$\hat{K}\psi(\vec{r}) + V(r)\psi(\vec{r}) = E\psi(\vec{r}) \tag{14.1}$$

First, remember that when we say $\psi(\vec{r})$, that's a shorthand for $\psi(x, y, z)$. On the left, we've broken the Hamiltonian $\hat{H}$ into the kinetic energy part $(\hat{K})$ and the potential. Notice that in here, we don't have $V(x, y, z)$, but only $V(r)$. When we say r without the little vector arrow, we mean the distance from the origin, that is, $r = \sqrt{x^2 + y^2 + z^2}$. The potential here is spherically symmetric. Because it depends only on the origin, if you rotated the whole system through any angle, the potential would be no different.

If we substitute in the correct expression for the potential $V(r)$, this equation becomes:

$$\hat{K}\psi(\vec{r}) - \frac{1}{4\pi\epsilon_0}\frac{e^2}{r}\psi(\vec{r}) = E\psi(\vec{r}) \tag{14.2}$$

As noted in the previous chapter, the potential is always negative. As r gets very large (i.e. the electron is very far away from the proton), the potential approaches zero. Again, this is just a convention; we could add any constant we wanted to the potential without changing the physics of what's happening. We've chosen this because it's convenient not to have to worry about the nuclei of atoms that are far away. The solution to this equation will be individual functions $\psi(\vec{r})$, each corresponding to a different allowed state, each with a corresponding energy level. The solutions that represent an atom— where the electron is bound to the proton— have $E < 0$. Classically, if a particle moves in this potential, that will set a maximum distance away from the origin that the particle could reach.

This page titled 14.1: The Schrödinger Equation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

14.2: The Orbitals

The specific functions $\psi(\vec{r})$ (or $\psi(x,y,z)$, or $\psi(r,\theta,\phi)$ that work with the Schrödinger equation are the eigenfunctions of the Hamiltonian operator. Because we don't know necessarily when we start what those functions are, even though we've now specified the operator, we have to go through a process in order to figure out what the possible functions $\psi(\vec{r})$ are.[1] Therefore, we call those functions the "solutions" of the Schrödinger equation. The exact functional form of these solutions aren't particularly illuminating. However, they do have some general features, the existence of which underlie all of atomic structure. In the previous chapter, we talked about some solutions that could be described by energy levels. For instance, the solutions to the one-dimensional simple harmonic oscillator have evenly spaced energy levels indexed by an integer n. In three dimensions, it will take three different "quantum numbers" to index the solutions. The numbers we choose to represent the states are the ones that have a most direct physical interpretation.

We call the solutions "orbitals" rather than just energy levels. In Hydrogen, there are usually multiple orbitals that share the same energy. In other atoms, fewer orbitals will share the same energy. These orbitals represent states available to the electron. Each state has an energy, a total angular momentum, and a z component of orbital angular momentum associated with it.[2] Those three values are the observables for which the orbitals are eigenstates. It's not surprising that the orbitals are energy eigenstates, because we produced them by finding states represented by wave functions that solve the Schr¨odinger equation, which is the energy eigenfunction equation. It is less obvious why they would be angular momentum eigenstates. It turns out that it's possible to break the kinetic energy operator into two parts, a "radial" part and a part that is the orbital angular momentum operator. That means that the Schrödinger equation includes the orbital angular momentum eigenvalue equation inside it.

These orbitals are not, however, eigenstates for position or for momentum. In particular, not being position eigenstates, electrons in atomic orbitals do not have definite position. Rather, there is a probability density for them to be at different positions, just as an electron in the $|+z\rangle$ spin state is not in a definite state of x spin and has probabilities to be found with positive and negative x spin. While the name "orbital" suggests that the electron is circling the nucleus in a manner analogous to how the Earth circles the Sun, this is not what is happening at all. The electron doesn't follow any particular path through space around the nucleus. Instead, the orbital is a probability cloud representing the effective amplitude for the electron to be found in any one tiny region of space around the nucleus. One consequence of this is that the negative charge associated with the electron is spread throughout this cloud. Whereas the Earth's mass is always at the position the Earth is in its orbit at any given moment, there is no single position for an electron in an atom, so there's no single place where the electron's charge is. The charge of the electron is more diffuse. Additionally, the center of the cloud is right at the nucleus. For the ground state, where the probability distribution is spherically symmetric, to something "far" from the atom (i.e. far enough away that the probability for the electron being found that far away or farther is negligible) the atom acts as if it were entirely neutral, with the effective charge of the electron being at exactly the same place as the effective charge of the proton.

Of the three quantum numbers that represent the electron orbitals, the quantum number n, the principle quantum number, is sometimes called the shell number. The average distance of the electron from the nucleus is determined primarily by n. In a single-electron atom, to first order the energy of the atom is determined entirely by n. (There are second order effects, such as the magnetic interaction between the spin and orbit of the electron, that are beyond the scope of this class.) The ground state has $n=1$, and higher shells have larger values of n.

The second quantum number, l, indexes the total orbital angular momentum of an electron in that state. It too must be a non-negative integer; that is, it can be 0 or a positive integer. The angular momentum represented by l does not include the spin angular momentum of the electron. Electrons remain electrons, and as such their total spin angular momentum is $\frac{\sqrt{3}}{2}\hbar$, and the possible projection along any axis is quantized to $+\hbar/2$ and $-\hbar/2$. The total angular momentum in the electron cloud of an atom depends on the orbital and spin angular momenta of all the electrons. Combining those angular momenta is fairly complicated, and involves taking into account the fact that electrons are indistinguishable particles. It turns out that for a given state, the angular momentum quantum number l must be less than the principle quantum number n. Thus, for the ground state, $n=1$, we know that $l=0$. The next shell out, $n=2$, there are two possible values of l: $l=0$ and $l=1$. The orbital angular momentum associated with a given value of l is $\hbar\sqrt{l(l+1)}$. Notice that this means that the ground state orbital has zero orbital angular momentum! This highlights the degree to which these orbitals are not analogous to planets circling stars, for a planet circling a star assuredly does have orbital angular momentum.

The third quantum number, m, indexes the z projection of the angular momentum. Just as electron spin angular momentum is quantized, so is electron orbit angular momentum. However, there's a difference. Instead of having half-integral values, orbital

angular momentum has integral values. As a spin-1/2 particle, we could say that the quantum number for the total spin of every electron is s = 1/2. The z projection, which we've called S_z, is then either $+\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$. We could say that there are quantum numbers, perhaps m_s, associated with electron spin that give the electron two possible spin states, one with $m_s = 1/2$, the other with $m_s = -1/2$.

In electron orbits, the total angular momentum l is an integer. m can be either positive or negative (representing angular momentum that's in the positive or negative z directions). Unlike electron spin, m can also be zero. However, m cannot get any larger than l. Thus, for a given value of l, there are $2l+1$ possible projections: $m = -1, m = -1+1, \ldots, m = 0, \ldots, m = l-1, m = l$.[3] The z component of orbital angular momentum associated with an orbital with quantum number m is just $m\hbar$. Notice that, just as with electron spin, it's impossible to have a z component of angular momentum that is equal to the total angular momentum of the state. Just like electron spin, the x, y, and z components of angular momentum are all represented by non-commuting operators. That is, an orbital can only be in a definite state for one of them. Thus, while a classical particle would have multiple different states (effectively) for a given l and m (as there would be different x and y projections of angular momentum available), a quantum particle's orbit is completely specified by just the total angular momentum quantum number l, and the z projection quantum number m.

The solution to the Schrödinger equation provides a dazzling wealth of orbitals available to the electron in the Hydrogen atom. Strictly speaking, we've solved for the orbitals as energy eigenstates. This means that if an electron is in one of those states, in the absence of observations or interactions it will stay in that state. (If it's in an indeterminate energy state, if you somehow manage to measure the energy of the atom you will collapse the atom's wave vector and it will drop into an eigenstate.) However, observationally, atoms that are in excited states do not stay there forever. After a while, they will spontaneously decay, with the electron dropping down to a lower state, and eventually with the electron reaching the ground state (which is stable. This would imply that the atom must somehow be interacting with something, if it is able to change from one eigenstate to another. Indeed, it does; it is interacting with the electromagnetic field. Even if there isn't any light (i.e. any "excitations of the electromagnetic field") around us, the field is always there. What's more, there are always virtual photons, as a result of the energy/time version of Heisenberg's Uncertainty Principle (Section 11.4.1). The interaction of the atom with the electromagnetic field yields a probability in any given time interval that the atom may emit a photon and drop to a lower energy state. The energy of that photon corresponds exactly to the difference in energy between the upper and lower states of the transition.

---

[1]That process is much more advanced than what could be covered here and involves differential equations.

[2]Remember that if a system is in a definite state for the z component of angular momentum, it can not have a definite value for x and y angular momentum. This is true for orbital angular momentum as well as for the intrinsic spin of particles. Whereas in a classical system, x, y, and z components of angular momentum would give you three "degrees of freedom", three things that could be varied, the orthogonality of those observables in quantum mechanics means that you only have two: total angular momentum, and one component.

[3]So for $l = 1$, there are three possible values of m: -1, 0, and 1. For $l = 2$, there are five possible values of m: -2, -1, 0, 1, and 2. And so forth.

This page titled 14.2: The Orbitals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

14.3: Visualizing Orbitals

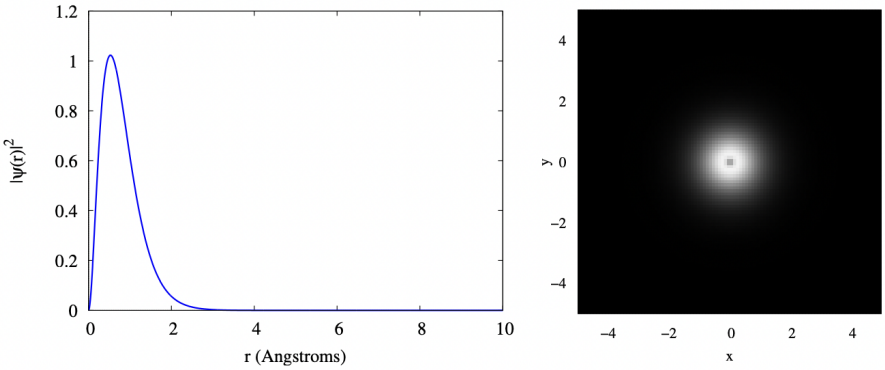
Historically, the energy available orbitals in an electron were identified by observing the wavelengths of light that would be emitted as the electron changed energy levels. Until we understood that light was quantized in photons, it was difficult to explain this as different energy levels. However, it was recognized that there were specific spectroscopic “lines” (i.e. wavelengths associated with a given transition) for each atom. The nomenclature that grew up around these lines is what gives us the nomenclature that we use today to name the various orbitals. Alas, the reasons for the letters that we use have little to do with the physics of the atom and everything to do with how the spectroscopic lines looked to the scientists who identified them. Each orbital is indicated by a form nx, where n is the principle quantum number and x is a letter corresponding to the angular momentum quantum number l. That letter is s for orbitals with $l = 0$, it is p for orbitals with $l = 1$; it is d for orbitals with $l = 2$; and it is f for orbitals with $l = 3$.

This page titled 14.3: Visualizing Orbitals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

14.3.1: s Orbitals

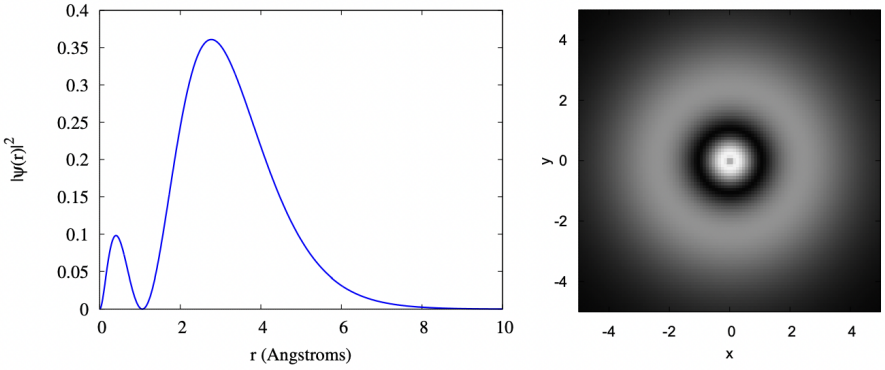
Orbits with $l = 0$ are called s orbitals. Although this is not where the letter comes from, it's useful to think of these as "spherical" orbitals, because they are spherically symmetric. However, they aren't just spheres! Again, remember that the probability cloud for the electron is a fuzzy ball around the nucleus, representing where the electron is likely to be found.

The plot below shows two visualizations of the 1s orbital. On the left is a plot of $\Psi^*(r)\Psi(r)$. This gives the probability density for the electron to be found at radius r. That is, you must pick a small range dr around the r you're interested in, and multiply this probability density by that dr. You then get the probability for finding the electron with that dr of your chosen r. On the right is a cut through the $x - z$ plane showing the probability density as a function of position. Lighter colors mean more probability of finding the electron at that position. Notice that there is a darker spot at the center. This corresponds to the probability dropping to zero at $r = 0$, as seen in the left plot. In both cases, distances are plotted in terms of Angstroms; one Angstrom is 10^{-10} m which, as you can see from the plot, is about the size of an atom.

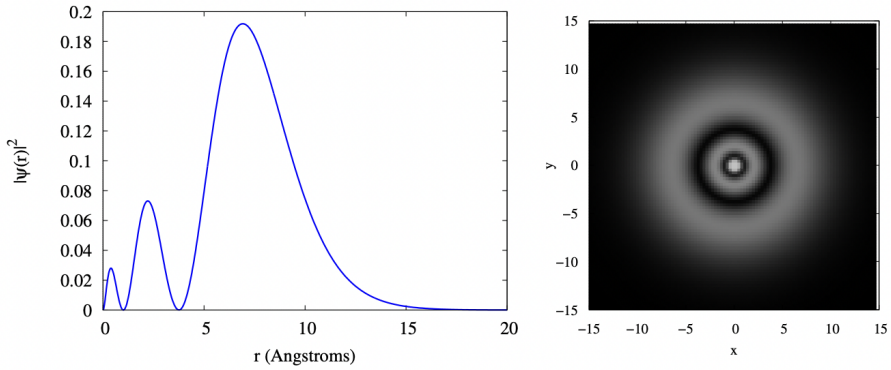


If you go to the 2s orbitals, an additional bump is added to the radial wave function. Also, the average distance the electron is from the center of the atom gets larger. While the probability clouds for a 1s and 2s orbital overlap, most of the probability for a 2s electron is outside most of the probability for a 1s electron. This means that to some extent, when working out the properties of an atom with two electrons in the 1s shell and one 2s electron (that would be Lithium), we can treat the nucleus plus the 1s shell as a single spherical ball of net charge +1. While this isn't perfect, this does lend some support to the approximation we'll make for multielectron atoms that each electron is moving in a nuclear potential and not interfering too much with other electrons.

Below are the same two plots for the 2s orbital. The scale of the axes is the same as the scale used previously in the 1s orbital, so that you may compare the plots directly.



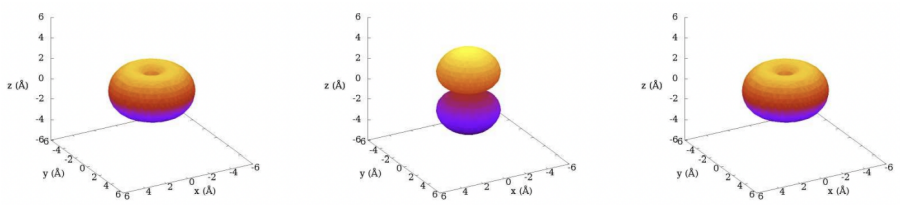
As we move to the 3s orbital, we have to expand the limits of our plots, as the electron is starting to have more and more probability to be at greater radius. In the plots below, you can see that the electron cloud still has reasonable probability density at a radius of 15 Angstroms. You can also see that the 3s orbital is three concentric fuzzy spherical shells; equivalently, the radial function has three bumps. Again, sizes are in Angstroms.



This page titled 14.3.1: s Orbitals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

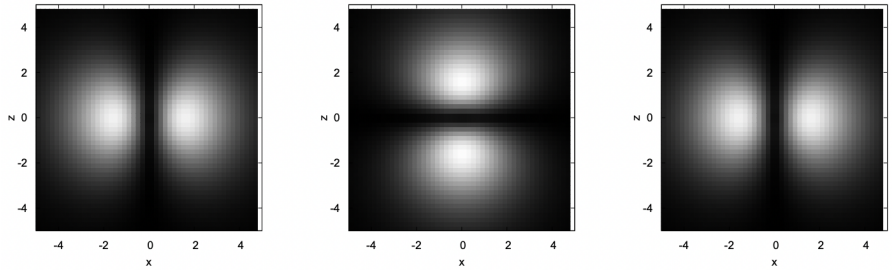
14.3.2: p Orbitals

Orbitals with total angular momentum quantum number $l = 1$ are called p orbitals. Remember that l must always be less than n. As such, the first shell has no p orbitals; it only has s orbitals. This means that you can only put two electrons (with opposite spin) in the first shell. In the second shell, you can put eight total electrons. You can put two electrons in the 2s orbital, and six in the 2p orbital. Why six? For $l = 1$, there are three possible values for m, the quantum number that indexes the z component of angular momentum: $m = 1$, $m = 0$, and $m = -1$. Below is three plots showing what the $2p$ orbitals look like.



$m = +1$ $m = 0$ $m = -1$

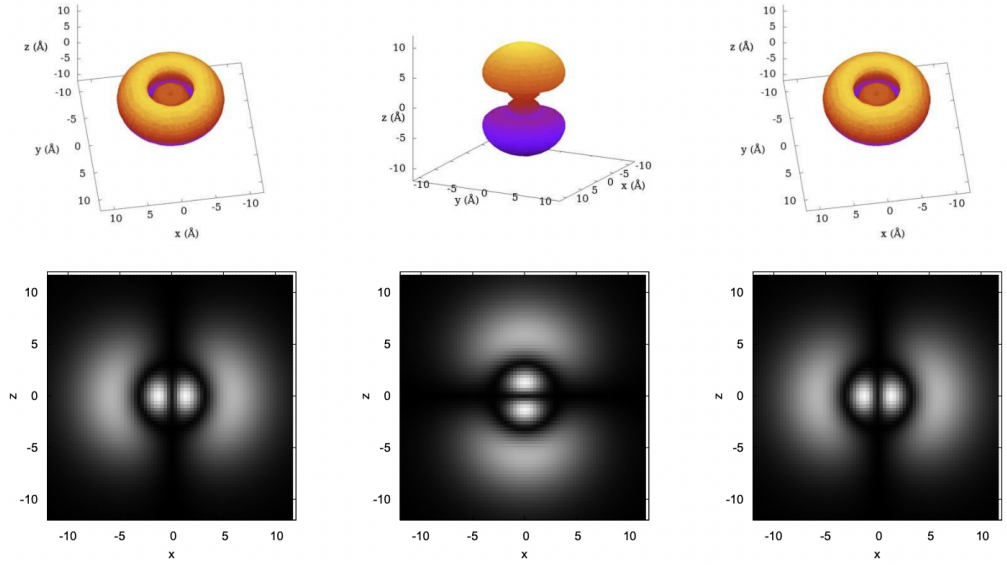
What we've drawn here is a surface of constant probability. In reality, the p-orbitals aren't hard shells, as this picture would seem to indicate. Rather, just as with the s-orbitals, they're fuzzy, with higher probability towards the "center" of the distribution which may not be at the origin!) and less probability away from it. As a way of visualizing this, the plots below show a cut in the $x - z$ plane of the three orbitals depicted above:



$m = +1$ $m = 0$ $m = -1$

Notice that the $m = +1$ and $m = -1$ $2p$ orbitals look identical. If you imagine rotating the plot around a vertical axis through the center of the plot, both of those orbitals look like two lobes, one over the other. The $m = 0$ orbital looks different, however. If you imagine rotating it around a vertical axis, you get an orbital that looks like a thick donut.

The p orbitals for higher values of n get more interesting. Just as the s orbitals become versions of themselves nested inside each other, the same thing happens with the p orbitals. Below are the three $3p$ orbitals:

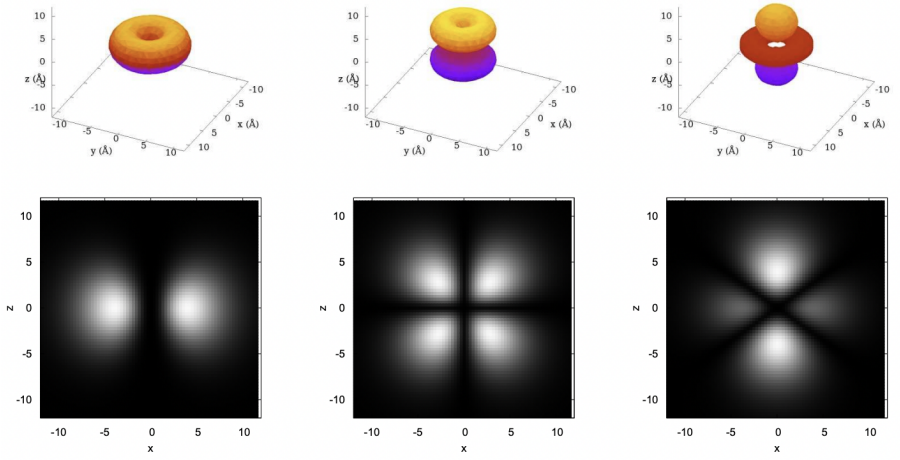


$m = +1$ $m = 0$ $m = -1$

This page titled 14.3.2: p Orbitals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

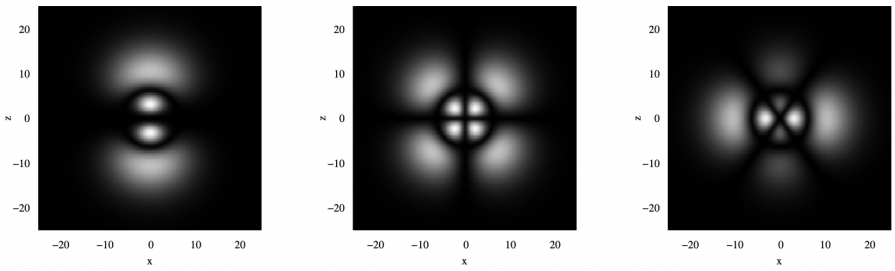
14.3.3: d Orbitals

When $l = 2$, the orbitals that are the solutions of the Hydrogen atom Schrödinger equation are called d orbitals. These orbitals only exist for shells with $n = 3$ and greater, again because l must be less than n. As we saw with the p orbitals, the probability density for the electron in space is the same for $+m$ and $-m$. As such, we'll only plot the positive-m versions of the orbitals. As before, in addition to a 3d plot showing "shells" at a constant probability level, there is a 2d plot showing a cut in the $x - z$ plane.



$m = 2$ $m = 1$ $m = 0$

As with the p orbitals, as we go to the d orbitals in higher shells they get more interesting. Plotted below are the cuts through the $x - z$ plane for the $4d$ orbitals:



$m = 2$ $m = 1$ $m = 0$

This page titled 14.3.3: d Orbitals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

15: The Periodic Table of the Elements

In Section 13.4, we saw the example of a square potential well. In that example, there were three bound states. This is the solution for a single electron moving in that potential, so each of these levels represents a possible energy state available to that electron. There are in fact six total states available, because there are two spin states available to an electron as well. Under the approximation that electrons do not interact with each other, we could take exactly these energy levels and put up to six electrons in the potential well. Because electrons are fermions, no two electrons can be in the same state. (If we were putting bosons into the potential well, there would be no limit, as you can put multiple bosons in one state.)

The process of constructing the periodic table of the elements is similar to the process of filling up this square well with electrons. In the previous chapter, we saw that the states available to an electron are indexed by three quantum numbers: n, the principle quantum number, l, the total orbital angular momentum quantum number, and m, the quantum number indexing the z projection of orbital angular momentum. In addition, there is electron spin, allowing two electrons to go into each $|n, l, m\rangle$ state.

On the periodic table, the "atomic number", usually indicated as the largest number in a display and often represented with the letter Z, is the total number of protons in the nucleus of the atom. The charge on the proton is exactly opposite the charge on the electron; whereas electrons have a charge of -1.602×10^{-19} C, protons have a charge of $+1.602 \times 10^{-19}$ C. Thus, for a neutral atom, the number of electrons is equal to the number of protons. Chemistry is all about the dynamics of electrons as atoms interact with each other, form bonds, trade electrons, and so forth. Therefore, from a chemical point of view, it might be more useful to think of Z as the number of electrons in a neutron atom of an element. (What if there is an additional electron added, making the atom negative, or if there is an electron removed, making the atom positive? In that case, we call it an ion, but we still name the ion based on the number of protons. A Chlorine atom with an extra electron would be called a negative Chlorine ion.)

Elements are constructed by filling in electron states until the number of electrons matches the number of protons in a nucleus. The number of states available at each shell is dictated by how angular momentum functions under quantum mechanics: as we saw in the previous chapter, l must be less than n, and m varies from $-l$ through l. If you put these two things together, it would be fair to say that angular momentum at the quantum level is responsible for the structure of the periodic table of the elements, the chemical properties of the different elements, and thus for chemistry and life as we know it. The number of states available is influenced, for instance, by the fact that angular momentum can only have definite states for projection along one axis at a time. The structure of the periodic table would be very different if x, y, and z angular momentum operators all commuted.

15.1: Interacting Electrons, Energy Levels, and Filled Shells

15.2: Filling Up Orbitals

15.3: Reading a Periodic Table

15.3.1: Electronic Configuration

This page titled 15: The Periodic Table of the Elements is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

15.1: Interacting Electrons, Energy Levels, and Filled Shells

In fact, electrons do interact with each other. In the previous chapter, we made arguments that these interactions should be smaller than the interaction with the nucleus. Because electron probability clouds are spread out, and outer shell clouds only have relatively small overlap with inner shell clouds, often, especially when viewing inner shells, you can approximate them as just lowering the net effective charge of the proton. That is, if you look at a Sodium atom, it has 11 electrons. The first 10 electrons will fill up the $1s$, $2s$, and $2p$ states. That leaves the outermost electron in the $3s$ state. Because there isn't a whole lot of probability for that $3s$ electron to be found where the inner electrons are usually found, you could approximate the situation for that outer electron that it's orbiting a ball of charge with a net charge of +1 (in atomic units), neglecting the fact that that charge is made up if +11 in the tiny nucleus and −10 in the outer electron cloud. However, even though interactions between electrons are secondary to the interaction between each electron and the nucleus, they are there, and they do ultimately have a lot of influence as to how elements at different places on the periodic table behave.

One of the primary effects of electron interactions is that the s, p, and d orbitals for a given value of n are not at exactly the same energy. In a Hydrogen atom—or any ion that only has one electron— they are, to a fairly good approximation. If there is more than one electron, however, the electron-electron interactions modify the energies of these states. In general, levels with higher l will be higher energy states than levels with lower l but the same n. In the absence of something external (such as a magnetic field), levels of different m but at the same n and l will still have approximately the same energy. Sometimes, you will find levels with a higher n but a lower l to be at a lower energy level than levels with a lower n and higher l. For instance, the $4p$ states tend to be filled before the $3d$ states. This isn't always a hard and fast rule; sometimes you will see the states filled out of the "standard" order. The interactions between electrons make the entire system a many-body system, and many-body systems are often notoriously difficult to solve in Physics.

For the most part, atoms are "happiest" (if you will allow for some anthropomorphization for purposes of discussion) if the number of electrons equals the number of protons. If there is one too many electrons, the ion will generally be happy to give away one of its negative electrons to the first positive charge that goes along. Likewise, if there is one too few electrons, the ion has an extra positive charge, and will tend to snap up any spare electrons in its vicinity.

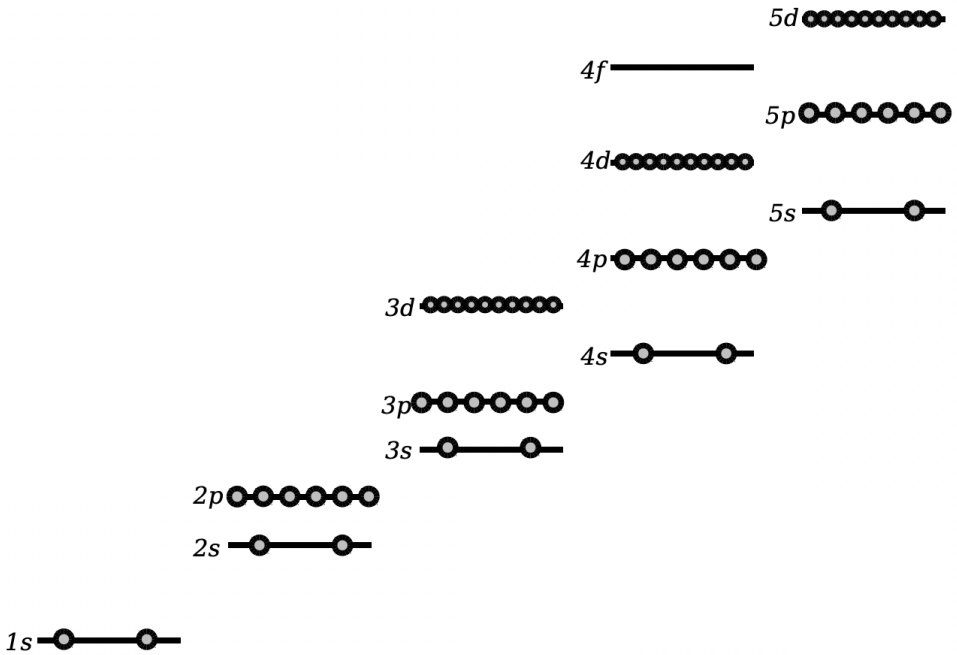
However, this is not the only consideration for atom happiness. Atoms also like to have a filled shell. That is, Helium is more chemically stable than Hydrogen, because whereas Hydrogen only has one of two possible electrons in the $1s$ state, Helium has entirely filled the $n = 1$ shell by placing two electrons in the 1s state. Likewise, Neon, with 10 electrons, has filled up both $1s$ states, both $2s$ states, and all six $2p$ states, making it a very chemically stable element. The elements down the right column of the Periodic Table are called "noble gasses". They are so called because they are chemically stable, and don't tend to interact with other atoms or form molecules. (They're noble, and thus above it all, or some such. Doubtless sociologists of science love to tear apart this nomenclature to display cultural bias in scientists.) The reason they are so stable is that each one of these noble gasses is an element that has just completely filled a set of p orbitals. (The one exception is Helium. It has completely filled the $n = 1$ shell, where there are no p orbitals.) Ne has completely filled its set of $2p$ orbitals. Ar has completely filled its set of $3p$ orbitals. Kr has completely filled its set of $4p$ orbitals. And so forth.

You can get a first guess at the chemical properties of an element by comparing how close it is to a noble gas. If an element has just one or two electrons more than a noble gas, the easiest way for it to be more like a noble gas would be for it to lose an extra electron. Elements like these are more apt to form positive ions than negative ions. An example is Sodium. Sodium has atomic number 11. The first 10 electrons fill up the $1s$, $2s$, and $3p$ orbitals; that is, they're like a Neon inner core. Then, just outside that, is a single $3s$ electron. If Sodium loses that electron, then it is electrically positive, but now it has a happy noble-gas-like electron configuration. In contrast, Chlorine has 2 electrons in the $3s$ shell and 5 electrons in the $3p$ shell. All it needs is one more electron to have a full $3p$ shell, giving it the electronic configuration of Krypton. If you put these two elements together, each Cl atom will tend to take away an electron from each Na atom, leaving the Cl a negative ion and the Na a positive ion. Those two ions then will have an electrostatic attraction towards each other as a result of their opposite charges. The result is a crystal, Sodium Chloride, more commonly known as salt. In this case, the bonds holding the crystal together are "ionic bonds". In most molecular bonds, an electron is shared between elements. In this case, however, the Sodium is so eager to get rid of an electron and the Chlorine is so greedy for another one that effectively the electron transfers all the way across from the Na to the Cl.

This page titled 15.1: Interacting Electrons, Energy Levels, and Filled Shells is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

15.2: Filling Up Orbitals

You can use the diagram below to figure out where an element will fall on the periodic table. Start with the lowest energy states, and fill in available states with electrons. Keep filling them in until you have as many electrons as you need. If you have just a single spare electron in an s shell, then that is an element like Hydrogen, Lithium, or Sodium; it goes on the left column of the periodic table. As you fill in more and more levels, you move farther and farther to the right on the table. If you completely fill a p orbital, then you're on the very right side of the table, and the element you've constructed is a noble gas.



States available to electrons in atoms. States higher in the diagram are (usually!) states at higher energy, although the exact spacing on the diagram should not be interpreted as meaningful. Although not drawn, there are 14 states available in the $4f$ orbital. Additionally, to fill out the current modern periodic table, we'd need to extend this to include the $5f, 6s, 6p, 7s$, and $7p$ orbitals.

Because the chemical properties of an element are approximately determined by how far away it is from a noble gas, constructing the periodic table in this manner means that every column should have somewhat similar chemical properties. Thus, each column on the periodic table is called a "Group". Each row on the periodic table is called a "Period". The name "Periodic Table" is not arbitrary. When something is periodic, it means that it has regular cycles. The period of the Earth's orbit around the Sun is about 365 days; after that much time, the Earth is back where it started. On the periodic table, it's element numbers that are periodic. After you add just enough electrons to fill up a p orbital, the next electron you add is as if you were starting over on a new shell.

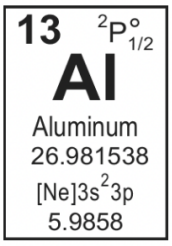
However, there is a difference with the Periodic Table. In a simple orbit such as the Earth going about the Sun, each period is exactly the same length. However, on the Periodic Table of the Elements, periods get longer and longer. As you go to higher and higher values of n, there are more and more states available. At a given n, you can have values of l between 0 and $n-1$. Thus, for $n=1$, there are only two states available, and thus the first period has only two elements: H and He. At $n=2$, there are now eight states available (two s states and six p states), so there are eight elements in the period: Li, Be, B, C, N, O, F, and Ne. You might then expect there to be 18 elements in the third period, as for $n=3$ there are 18 states: two s states, six p states, and ten d states. However, it turns out that the $4s$ states are at a lower energy level than the $3d$ states. Thus, the third period only fills up the $3s$ and $3d$ states, and has eight elements just like the second period. In the fourth period, starting with Potassium, we fill in the $4d, 3d$, and $4p$ states (approximately in that order), and now have eighteen elements. The same thing happens with f orbitals; it isn't until the sixth period, after the $6s$ states are filled, that the $4f$ orbitals start to get filled.

The fact that as you go to higher and higher periods, there are more and more states available before you can completely fill a p orbital, is what gives the periodic table its iconic "stepped" structure. The first step comes after Hydrogen and Helium. The first period only has the $1s$ orbital available, and only has two elements in it. The second and third periods each have 8 elements, filling up the $2s, 2p, 3s$, and $3p$ orbitals. The fourth period now has 18 elements in it, because in addition to the $4s$ and $4p$ orbitals, it also has to fill up the $3d$ orbitals.

This page titled 15.2: Filling Up Orbitals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

15.3: Reading a Periodic Table

If you look at a periodic table, there is a variety of information you may find on it. Every periodic table includes the symbol of the element (one or two big bold letters at the center of the element's box), and the atomic number of the element (the number of protons in that element, usually shown as a number in the upper left). For example, below is the entry for Aluminum from the NIST Periodic Table (Dragoset et al., 2003).



Here, you can see the symbol for the element is Al, and the atomic number is 13.

Usually, below the element symbol, you will find the atomic weight of the element. This is in units of "atomic mass units" or amu; one amu is equal to 1.66×10^{-27} kg. In this example, the atomic weight is given to eight significant figures, and is 26.981538 amu. The atomic weight in amu is approximately the number of protons and neutrons— which, together are just called "nucleons"— in the element, but there are a number of complications. First is the complication that for some elements, there are multiple isotopes. Different isotopes have the same number of protons but different numbers of neutrons. For example, the atomic weight of Chlorine is 35.45 amu. This is largely because in nature, we find Chlorine in two isotopes: Cl-35 and Cl-37, with 35 and 37 total nucleons respectively. There is a second complication, however. The mass of a nucleus is not exactly equal to the sum of the masses of the protons and neutrons that compose it! Each nucleus has what's called a binding energy. This binding energy is equivalent to the 13.6eV of energy that holds an electron on to a Hydrogen atom. It is the total energy for all of the nucleons in their bound states, and is negative for a stable nucleus. This binding energy is taken away from the effective mass of the nucleus, using the conversion $E = mc^2$. In fact, exactly the same thing is true for atoms! However, the binding energy compared to the mass of atoms is something like one part in a billion, so as such when dealing with chemical reactions and other electronic transitions, we can approximate mass as being conserved. Nuclear binding energy can get up to a few percent of the total mass. (It is this difference that makes nuclear power so much more efficient, in terms of energy produced per mass of fuel used, than chemical power.)

This page titled 15.3: Reading a Periodic Table is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

15.3.1: Electronic Configuration

Often, but not always, the periodic table will include one or two sets of symbols intended to convey information about the ground state of the atom's electron cloud. The one seen more rarely includes information about the spin, orbital, and total angular momentum of the atom. In the NIST entry for Aluminum above, you can see this information in the upper right as $^2\mathrm{P}^{\mathrm{O}}_{1/2}$. The letter in the middle represents the total orbital angular momentum of all of the electrons the atom put together; S means $l=0$, P means $l=1$, D means $l=2$, and F means $l=3$. For Hydrogen, the single electron is in the $1s$ orbital, which has no orbital angular momentum, so that letter is S. For Aluminum, the letter is P. This results from the single electron in the $3p$ state. There are two electrons in each of the $1s, 2s$, and $3s$ states, and none of them have any orbital angular momentum. However, for Aluminum, there are also six electrons in the $2p$ state. But, because that shell is filled, there will be as many electrons with z-angular momentum of $+\hbar$ as there are with $-\hbar$ (two each, in this case), so all of their orbital angular momentum cancels out.

The superscripted number before the letter tells you about the electron spin state. It is equal to $2s+1$, where s is the net electron spin divided by $\hbar$. For Hydrogen, this is $^2S_{1/2}$. In Hydrogen, the net electron spin is 1/2, because there is just one electron that has spin $+\hbar/2$, so $2s=1$. In Helium, that number is 1, because the two electrons have spins in the opposite directions, so $s=0$. For aluminum, the net electron spin is the result of a single electron in the $2p$ state (as all of the filled states will have as many spin up as spin down electrons, thereby cancelling out each other's angular momentum). Thus, the net electron spin is 1/2 (as always in units of $\hbar$), so the number we see in the example above is (2)(1/2) + 1 = 2. Finally, the subscripted number after the letter is J, the quantum number associated with the total electronic angular momentum of the atom. J represents a combination of orbital and spin angular momentum for the electrons. For Hydrogen, $J=1/2$, because the angular momentum is entirely in the spin of the electron; for Aluminum, it's also $J=1/2$, but the reason is more complicated. Both the spin and orbit of the $3p$ electron contribute, but it would have been possible for them to combine yielding either $J=1/2$ or $J=3/2$. For Helium $J=0$ because there is no net angular momentum: there's no orbital angular momentum for two electrons in the $1s$ state, and the two spins cancel each other out. (As a caution, adding angular momenta in quantum mechanics can become complicated for cases with higher numbers than these examples.[1]) Similar to what we see with orbital angular momentum, the physical amount of angular momentum for an atom with total orbital quantum number J is $\hbar\sqrt{J(J+1)}$.

(The superscripted O— it's a capital O, not a zero— on the notation you see for Aluminum indicates that Aluminum has "odd parity". Parity is another quantum property that you don't need to worry about here.)

You may also see a series of letters and numbers that tell you how many electrons there are in each orbital. For Hydrogen, this is $1s^1$. That is, there is but a single electron in the $1s$ orbital. For Helium, it's $1s^2$: there are two electrons in the $1s$ orbital. For Nitrogen, it's $1s^2 2s^2 2p^3$. The $1s$ and $1s$ orbitals are filled, and the $2p$ orbital is half-filled. For periods after the second, it's often conventional not to list the full state, but to list the noble gas that has the same configuration as the inner core of electrons, and then just the states of the electrons outside of that. In our example of Aluminum above, the configuration is $[\mathrm{Ne}]3s^2 3p$ It's got all the electrons that Neon does— $1s^2 2s^2 2p^6$ — plus an additional two electrons in the $3s$ state and one in the 3p state.

The number at the very bottom of the box for Aluminum, 5.9858, is the ionization potential in eV for Aluminum. Many periodic tables will not include this number. This is the amount of energy it takes to remove one electron from the atom, forming a positive ion.

---

[1]For example, with Aluminum, it turns out that for the total spin+orbit angular momentum of the $3p$ electron to be in a definite state $J=1/2$, neither the orbital z component nor the spin z component may individually be in definite states. If we write the state of the $3p$ electron as $|m, s_z\rangle$, for an Aluminum atom to have $J=1/2$ and $J_z=1/2$ (i.e. we've specified the orientation as well as the total angular momentum of our Aluminum atom), then the angular momentum state of the outermost electron would be

$$\sqrt{\tfrac{2}{3}}|+1,-1/2\rangle - \sqrt{\tfrac{1}{3}}|0,+1/2\rangle$$

This page titled 15.3.1: Electronic Configuration is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

CHAPTER OVERVIEW

16: Matter

We've concentrated primarily on electrons throughout this course. Indeed, in our everyday life, it is the interactions of electrons that, together with photons (light), drive most of what we do. In this final chapter, we'll peer down inside the atom to see what the most fundamental particles are, and then extend our view out to bulk states comprised of large numbers of electrons.

16.1: The Standard Model of Particle Physics
16.2: Nuclei and Atoms
16.3: Molecules
16.4: Solids
16.5: Liquids, Gasses, and Plasmas
16.5.1: Quantum Gasses
16.6: Planets, Stars, Galaxies, and Clusters
16.7: Dark Matter and Dark Energy

This page titled 16: Matter is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

16.1: The Standard Model of Particle Physics

Our best current understanding of Physics at the most basic level is that it is composed of a number of fundamental particles. These particles are, as best we can tell, points, much like electrons (which are in fact one of the fundamental particles). They have various properties associated with them, including mass, spin (angular momentum), electric charge, and others. It is from these fundamental particles that all of the matter we interact with is built. (However, matter built from these particles only makes up 5% of the total density of the Universe! See Section 16.7.)

Broadly, we can divide the particles into two categories: fermions and bosons. Matter is built from fermions, and the interactions between matter— the four forces— are carried by bosons.

The fermions are divided into three generations; it is only the first generation that we ever interact with on a daily basis. Each generation includes two quarks and two leptons. In the first generation, the quarks are called the "up" and "down" quarks. The two leptons are the familiar electron, and the electron neutrino. Each of the other two generations has an additional two quarks, heavier than the quarks in the first generation; and an additional two leptons, including a heavier version of the electron, and a corresponding neutrino. Additionally, for each fermion, there is a corresponding antiparticle. For example, the anti-electron is called the "positron". Antimatter is rare. In the very early Universe, we believe that matter and antimatter were present in equal quantities, but for some reason that is not understood matter had a slight advantage. A particle of antimatter, when it meets its corresponding matter particle, will mutually annihilate both, releasing their mass as energy according to the conversion $E = mc^2$. That slight advantage that matter had in the very early Universe is what's left over to build up stars, galaxies, and us today.

Table of fermions from the standard model of particle physics. All data is from the PDG (Nakamura and Particle Data Group, 2010). Charges are in units of e, the elementary charge. *: There are three neutrinos. Although masses are not known, we have limits that the heaviest one is less than 18.2 MeV, the middle one is less than 0.19 MeV, and the lightest one is less than 1.3 eV. However, the mass eigenstates and the flavor (i.e. "type of neutrino") eigenstates of the neutrinos are not the same, so it's impossible to identify a given mass with a given type of neutrino.

Generation	Fermion	Symbol	Charge	Spin	mc^2
First	Down Quark	d	-1/3	1/2	5.05 MeV
	Up Quark	u	+2/3	1/2	2.49 MeV
	Electron	e	−1	1/2	0.511 MeV
	Electron Neutrino	ν_e	0	1/2	*
Second	Strange Quark	s	-1/3	1/2	101 MeV
	Charm Quark	c	+2/3	1/2	1.27 GeV
	Muon	μ	−1	1/2	106 MeV
	Muon Neutrino	ν_μ	0	1/2	*
Third	Bottom Quark	b	-1/3	1/2	4 GeV
	Top Quark	t	+2/3	1/2	172 GeV
	Tauon	τ	−1	1/2	1.78 GeV
	Tau Neutrino	ν_τ	0	1/2	*

Quarks are never observed in isolation. For the most part, they are observed in bound states called hadrons. The proton and neutron are two three-quark particles. There is a huge zoo of additional hadronic particles, including baryons (made of three quarks) and mesons (made up of a quark and an antiquark). The proton and the neutron are the only stable hadrons. Indeed, even the neutron is not stable unless it's bound into an atom; a free neutron will decay to a proton, an electron, and an antineutrino in about 15 minutes. "Virtual" mesons are found inside the nucleus. Transient mesons and other sorts of baryons are made in particle accelerators, and also when cosmic rays hit the Earth's atmosphere.

In addition to the fermions that make up matter, there are four forces through which matter interacts. The most familiar of these forces are gravity and the electromagnetic force. In our current theories of physics, gravity is described by General Relativity, and is not included in the Standard Model of Particle Physics. We believe that we will one day be able to produce a working theory of quantum gravity, but we have yet to successfully do that. We expect this theory to include the graviton as a massless spin-2 boson.

The electromagnetic force is the best understood of the four forces. It unifies the electrostatic force and the magnetic force. The photon is the particle that carries the electromagnetic force; we see it as light. Radio waves, infrared radiation, ultraviolet radiation, x-rays, and gamma rays are all forms of light at wavelengths different from those our eye can detect. All of these are made up of

photons, the quanta of the electromagnetic field. It is the electromagnetic force that forms the potential in which electrons move in atoms, and it is that interaction that governs the interactions between atoms.

Because the photon and the (presumed) graviton are massless, both gravity and electromagnetism are long-range forces. In contrast, the other two forces are short ranged. The strong nuclear force is the force that binds quarks together into protons and neutrons, and that ultimately binds protons and neutrons together into nuclei. The bosons that carry the strong nuclear force are called gluons. They may be massless, although a moderate mass isn't ruled out. However, other properties of the strong force limit it to a short-range force. The weak nuclear force is, as its name suggests, much weaker than the strong nuclear force, and has only a secondary effect in nuclei. The charge carriers of the weak force are indeed massive, limiting it to a short range force. The weak force is the only force other than gravity that interacts with all of the particles in the standard model. Neutrinos, in particular, only interact via the weak force, making them extremely hard to detect. The weak force is responsible for much radioactive decay; it is a result of the weak force, for instance, that a free neutron will decay into a proton, an electron, and an antineutrino. The bosons that carry the weak force are called "intermediate vector bosons." Their name is the same as their symbol. There are three, the W^+, W^-, and Z^0 bosons.

Table of bosons in the standard model of particle physics. All charges are in units of e, the elementary charge. Not included is the hypothesized graviton, which would be a massless, chargeless, spin-2 boson.

Force	Boson	Symbol	Charge	Spin	Mass
Electromagnetic	Photon	γ	0	1	0
Strong	Gluon	g	0	1	0?
Weak	W^+ W^- Z Boson	W^+ W^- $\mathbf{Z^0}$	+1 −1 0	1 1 1	80 GeV 80 GeV 91 GeV

This page titled 16.1: The Standard Model of Particle Physics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

16.2: Nuclei and Atoms

Quarks bind together to make protons and neutrons. A proton is composed of two up quarks and a down quark, and a neutron is composed of two down quarks and an up quark. Together, protons and neutrons are called nucleons. Nucleons can themselves bind together to make nuclei. These nuclei are always positively charged, with the total charge depending on the total number of protons. Nuclei are so called because they sit at the nucleus of atoms; an atom is a nucleus that has gathered negative electrons into the various electron states (i.e. orbitals) allowed by the electric potential created by its positive charge. The smallest nucleus is the Hydrogen nucleus, composed of but a single proton; it is about 10^{-15} m across. The largest stable nuclei have a more than 200 nucleons in them. Nuclei with more nucleons than that tend to be unstable, and spontaneously fission into smaller nuclei.

The binding energy— that is, the potential energy that results from combining nucleons together to make nuclei— is an appreciable fraction (1% or so) of the mass energy of the nucleons. The nucleus with the greatest binding energy per nucleon is Iron-56, which makes it (in a sense) the most stable nucleus. You can get energy out by fusing lighter elements together until you get to Iron-56; after that, it costs energy to build up heavier elements.

Of course, in nature, most materials are mostly electrically neutral, at least on Earth. It turns out that most of the baryonic material in space is in the form of plasma (mostly ionized Hydrogen, i.e. free protons and free electrons) filling the void between galaxies inside galaxy clusters. On Earth, though, for the most part if there's a free electron, it will be captured by the first nucleus that comes by with an extra positive charge. Hence, in our everyday experience, all things are composed of atoms. We organize our understanding of the various different types of atoms via the Periodic Table, as was discussed at length in Chapter 15.

This page titled 16.2: Nuclei and Atoms is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

16.3: Molecules

Atoms can bond together. Sometimes, if one atom is able to completely steal an electron from another atom (as is the case with Chlorine and Sodium atoms, where a Sodium atom will donate an electron to a Chlorine atom), the resulting ions will then stick together as a result of the electrostatic attraction between their opposite net charge.

More common, however, are molecules made from what is called covalent bonds. The electrons in the outer (unfilled) shell of an atom are known as “valence” electrons. Depending on the electronic configuration of an atom, it will have one or more effective valence electrons. In a molecule, the valence electrons are no longer associated with a single atom, but instead are shared between the electrons. In terms of the quantum mechanics involved, you wouldn’t find a solution to Schrödinger’s Equation for just the potential of one atom. Rather, you create a joint potential for the two atoms (including the effects of inner-shell electrons), and determine a solution for the system as a whole. The result is an electron wave function that indicates the electron probability cloud is shared between two or (for more complicated molecules) more of the atoms that composes the molecule. Just as nuclei have a binding energy, molecules have a binding energy, meaning that it is a lower energy state for these atoms to bind together and share an electron than it is for them to stay separate. Although this binding energy is typically a billionth of the mass energy of atoms, it is enough to create the vast majority of energy producing processes (e.g. burning gas to power a car) that we are familiar with in our everyday lives.

Finding these solutions to multi-atom potentials is an extremely difficult problem, and cannot be solved analytically (as the Hydrogen atom may be). Describing the quantum mechanical state and electron orbitals of any molecule more complicated than something like H_2 generally involves both heavy-duty numerical calculations on computers and heavy-duty quantum chemists.

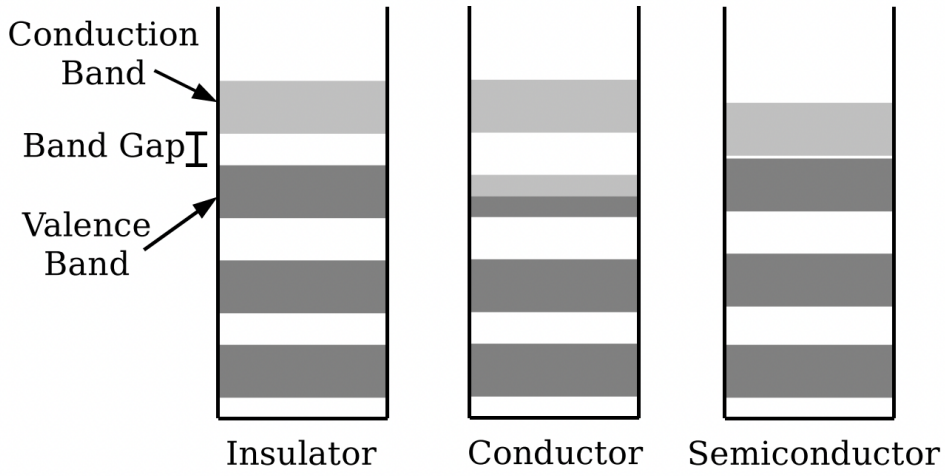
This page titled 16.3: Molecules is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

16.4: Solids

Roughly speaking, a solid is when a large collection of molecules are held together and fixed in place. They aren't completely still, unless a solid is at absolute zero temperature. (And that's not possible, as a result of Heisenberg's Uncertainty Principle.) Most solids are at a higher temperature; the solids you interact with every day are approximately at room temperature, about 20^{o}C or 290 K. At such a temperature, the molecules are vibrating about, each with about 1/40 eV of kinetic energy in that vibration. Solids are held together different ways. Some solids form crystals, where every atom is bonded one way or another to neighboring atoms. Other solids do not strictly have covalent or ionic bonds holding everything together, but are held together by intermolecular forces resulting from the attraction of one electron cloud to a neighboring nucleus and the like.

In some particularly interesting solids, valence electrons don't end up just being shared between neighboring atoms in molecules, but are rather associated with the solid as a whole. These solids may be conductors, if there are quantum states available for electrons to move about. When the valence electrons are associated with the solid as a whole, you can approximate the potential in which they move as as threedimensional square well the size of the solid. Of course, this isn't exactly true, as there will be localized potential wells where each atom in the solid is. However, it is reasonable to approximate the solid as a fixed lattice, with valence electrons potentially free to move through it.

If you associate the valence electrons with the solid as a whole, you then need to think about the states available to those electrons. Since we are talking about a chunk of material that may have a huge number of atoms (e.g. about 10^{22} atoms if we're talking a mere 1g sample of Copper), we're also talking about a huge number of valence electrons, and a huge number of available states. Depending on how you model it, you can think of the states as resulting from the lattice, or as resulting from the effective square well potential in which the electrons move. The nature of the underlying lattice does matter. Typically, the states available to the valence electrons come in bands of energy, with gaps between the bands; this is one way in which a simple square well does not reflect the nature of the potential (where the distribution of states would be continuous). Three examples of solids with bands and band gaps are shown below. Each diagram is an energy level diagram. In a shaded band, there are many electron states stacked one on top of the other. Because electrons are fermions, however, each individual state may only have two electrons (two, not one, because of electron spin).



Electron energy states in the band gap model of solids. Increasing energy is upwards on all three diagrams. Dark gray bands indicate bands of many energy states that are filled. Light Gray bands indicate bands of many empty energy states.

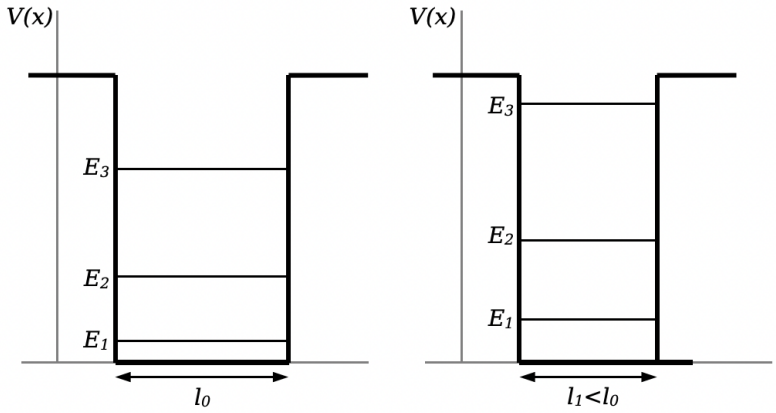
In the figure on the left, the top filled band is called the valence band. Again, this band, and the band below it, represent a huge number of closely spaced energy states for valence electrons. The energy of the top filled state (assuming the solid as a whole is in its lowest overall energy state) is called the Fermi energy. These states are not places where the electron can be, in the sense of places in space. Electrons in higher states have more energy than electrons in lower states, so they are in fact moving around. However, they're not moving around in a way that would allow electrons to flow in a coordinated fashion from one side of the solid to another. Rather, they're moving around in the same way that an electron in a higher orbital of an atom that has some kinetic energy associated with it. The electron is still tied to the solid, and stuck in the energy state it's in. But, that state is not localized to one nucleus; it's associated with the insulator as a whole. That there are no empty states nearby is why nothing can change that would allow for electric conduction.

The first empty band above the filled band is the conduction band. If you want to conduct electricity— that is, allow charge to move through the solid— you need to be able to get electrons out of the states where they're fixed, and into states that have empty states nearby. If there is a band gap between the valence band and the conduction band, this is not very easy; it takes a lot of energy to get an electron out of the state where it normally is and into a state where there are many neighboring states, allowing it to move in a coordinated way through the solid. The conduction band is the band where there are lots of empty states that the electron can make transitions between (or occupy in a superposition of states) so as to allow it to move through the solid and thereby carry electric current. (Electric current ultimately is just the transfer of net electron charge from one side of the conductor to the other.) Every material will ultimately carry electron current, if you apply enough potential to it. However, the wider the band gap, the more energy it takes to liberate an electron from the valence band and into the conduction band, thereby allowing it to move freely. If there is a wide gap, we would consider the material an insulator.

In the middle figure above, the material is a conductor. Here, the Fermi energy is in the middle of a band of available energy states. As such, electrons in the top filled states have many nearby states available to them. They may freely change states, and thus are able to go into states that can correspond to electrons transferring charge through the material.

The rightmost figure above is a semiconductor. Here, there is a band gap between the valence and conduction bands, but that band gap is quite small. Because of thermal excitations— electrons bumped about by the vibrating atoms resulting from the non-zero temperature— a tiny fraction of the electrons will in fact be up in the conduction band. However, the conductivity of this material remains small, because not very many electrons are up there. (The conductivity will go up with temperature, however, as more electrons statistically get bumped up into the conduction band.) The typical band gap in a semiconductor is around 1–2 eV, which is a lot larger than the 1/40 eV that is the average energy available to one particle at room temperature. By doping a semiconductor— that is, adding impurities— you can either add additional valence electrons inside the band gap at an energy just below the conduction band, or additional valence states at an energy inside the band gap just above the valence band. By putting such doped semiconductors together, you can create devices with interesting electrical properties, such as diodes and transistors.

One thing that is interesting about solids is what happens when you try to compress the solid. In a square well, if you decrease the width of the well the energy of all the states increases:



If you think about squeezing a solid, what this means is that you're pushing the states occupied by the valence electrons to higher energy levels, and therefore you must be putting energy into the solid. The need to exert energy on a solid when squeezing will manifest as a pressure (force per area) that resists the force trying to squeeze the solid. While you might think that it is the atoms and molecules themselves resisting being pushed closer together that makes a solid resist being compressed, this is not the whole case. In fact, this additional energy put into the valence electrons— which, remember, are not associated with individual electrons, but with the solid as a whole— contributes significantly to the restorative pressure of a compressed solid. This pressure is called Fermi degeneracy pressure. In this circumstance, "degeneracy" is a technical term referring to electrons all being packed into states as tightly as they can. The Fermi degeneracy pressure that resists the compression of a solid is a direct result of those states rising to higher energies as a result of the compression.

This page titled 16.4: Solids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

16.5: Liquids, Gasses, and Plasmas

If you put enough energy into a solid, eventually it will melt. At this point, the atoms and molecules in the solid are no longer held together in the crystal, lattice, or other structure. Instead, they have enough energy to break whatever bonds (covalent, ionic, or otherwise) holding them together, and now they can flow past each other. The atoms and molecules are still largely packed together as closely as they can go, and there still are bonds of a sort holding the broadly together, but no longer are they fixed in place. Such a state would be called a liquid.

If you raise the temperature of a liquid enough (to the “boiling point”), and continue to add energy, you can break the residual forces holding the liquid together, and give each molecule enough energy that it starts to bounce about freely. Add enough energy, and the liquid becomes a gas. At this point, each molecule or atom of the gas moves around freely. Molecules do very regularly collide with each other, but they’re not in constant interaction any more as is the case with a liquid.

Indeed, if you investigate the momentum and kinetic energy states available to molecules in a gas, the states are mostly empty. Unlike the valence electrons in a solid, which mostly fill up the lowest states available to them, there are many, many empty lower energy states for every gas molecule. As such, it’s easy to change the energy of a gas molecule by a very small fraction of its current energy, as there are so many empty states about. At this point, you can approximate the available energy states as a continuum, and the gas molecules behave like classical particles. A traditional gas is described by the Maxwell Boltzmann distribution, which specifies the fraction of gas molecules that will be moving at any velocity given the temperature of the gas and the mass of each molecule. This description applies to the air around us, to the gas between the stars, and to the gas in atmospheres of stars.

Solid, liquid, and gas are the standard “three states of matter.” If you ionize a gas— that is, if you tear an electron off of a substantial fraction of the gas atoms— it becomes a plasma, the fourth state of matter. Because the individual particles in a plasma are electrically charged (either positive ions or negative electrons), electric and magnetic fields can greatly influence the behavior of a plasma. There are a few ways to create a plasma. One is just to raise the temperature of the gas high enough so that the average kinetic energy of any particle is high enough that collisions will tend to ionize gas molecules. Another is to shine ionizing radiation — generally ultraviolet or x-ray radiation— on the gas. Interstellar gas around young massive stars is typically mostly ionized as a result of the radiation from those stars, even though the temperature of the gas itself isn’t high enough to maintain that ionization. A final way to ionize gas is to run high energy particles through it. For instance, if you can shoot an electron beam through dilute gas, it will tend to ionize the gas it passes through. This is how plasma discharge tubes are created.

This page titled 16.5: Liquids, Gasses, and Plasmas is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

16.5.1: Quantum Gasses

It is possible to create a fundamentally quantum gas, however. If you can lower the temperature of a gas enough while allowing it to stay as a gas e.g. by keeping it at a low enough density that it does not condense into a liquid or solidify), you can get to the point that a substantial fraction of the gas is occupying the lowest states available to it. At this point, the gas is no longer adequately described by classical physics. If the gas is composed of fermions (i.e. each molecule has net half-integral spin), you will have what's called a "Fermi gas", that is analogous in many ways to the valence electrons in a solid. If, on the other hand, the gas is composed of bosons, and you can lower its temperature enough, it's possible to create a "Bos-Einstein condensate", where a substantial fraction of the gas molecules all drop into the same state (something that would be impossible for atoms). At this point, you can see coherent quantum phenomena for the whole gas, such as interference, because of all the molecules or atoms that are in the same state. A Bos-Einstein condensate was first created in 1995 (Anderson et al., 1995); this work received the Nobel Prize in physics in 2009.

This page titled 16.5.1: Quantum Gasses is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

16.6: Planets, Stars, Galaxies, and Clusters

Once you get past the sizes of everyday solids, liquids and gasses, you enter the realm of astronomical objects. In our Solar System, such objects range from lowly asteroids, through dwarf planets such as Pluto or Ceres, through rocky planets such as the Earth or Mars, on up through the gas giants such as Saturn or Jupiter. However, the vast majority of the mass of our Solar System is in the Sun, the star about which everything else orbits. The Sun is a ball of gas, 300,000 times the mass of the Earth. It is composed of about 74% Hydrogen, about 24% Helium, and 2% everything else. This is a very different composition from the Earth, which is composed mostly of heavier elements. However, the composition of the Sun is more typical of the composition of the Universe as a whole— indeed, the Sun has a greater than average fraction of heavy elements!

Stars collect together into galaxies, gravitationally bound systems of millions, billions, or even trillions of stars. Galaxies themselves collect together into groups and clusters, which may themselves have thousands of galaxies. The groups and clusters we can identify are part of a filamentary structure that fills the Universe. Most galaxy groups and clusters are found in this filamentary web, with vast voids between them. On the largest scales, the universe is homogeneous. That is, if you look at one spot in the universe about a billion or so light-years across, it looks pretty much the same as another, with galaxy clusters on filaments surrounding voids.

This page titled 16.6: Planets, Stars, Galaxies, and Clusters is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

16.7: Dark Matter and Dark Energy

All of the matter discussed so far in this chapter only makes up 5% of the energy density of the Universe. (We can talk about mass density and energy density interchangeably, as mass is a form of energy. The amount of energy E in mass m can be found through the famous conversion $E = mc^2$.) Several different lines of evidence have shown us that most of the mass in galaxy clusters is not the luminous mass we can observe. Indeed, this mass can't be baryonic at all. Dark Matter is the name given to this mass; it's not dark the way dust is, which absorbs light is. It neither absorbs nor emits light; it only interacts with light gravitationally (and has been observed partly through the gravitational lensing effect). Dark Matter interacts only via gravity and, perhaps, the Weak Force. In this way, it is similar to neutrinos. We have not identified a particle that can make up Dark Matter. We're very sure that it's there, and we're very sure that it makes up most of galaxies and galaxy clusters, but we don't know exactly what it is.

Dark Matter, however, itself only makes up about 20-25% of the energy density of the Universe. In the late 1990's, astronomers discovered that the expansion of the Universe is accelerating; this discovery was awared the Nobel Prize in Physics in 2011. This is not something you would expect from regular matter or from Dark Matter. With matter (including Dark Matter), as well as normal forms of energy such as radiation (light), the gravitational effect is attractive. The result would be to tend to slow down the expansion of the Universe. For the Universe to be accelerating, there must be something else in it. Dark Energy is the name given to this unknown substance that evidently makes up about 75% of the energy density of the Universe and that has a negative gravitational effect. Dark Energy is even more unknown than Dark Matter, and indeed some believe that it doesn't really exist as a substance, but is a pointer to our theories of gravity breaking down. Many believe that the most likely candidate for Dark Energy is vacuum energy (see Section 11.4.2), but for now nature of Dark Energy remains one of the primary outstanding unanswered questions in both astronomy and fundamental physics.

This page titled 16.7: Dark Matter and Dark Energy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Pieter Kok via source content that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

Index

A

annihilation operators
8.2: Creation and Annihilation Operators

B

Bloch sphere
3: Schrödinger and Heisenberg Pictures
Born rule
4.3: Imperfect Measurements

C

coherence
4.2: Decoherence
Copenhagen interpretation
2: The Postulates of Quantum Mechanics
Coulomb integral
9.1: Interacting Electrons in Atomic Shells
creation operators
8.2: Creation and Annihilation Operators

D

de Broglie wavelength
13.3: Free Particles and the de Broglie Wavelength
decoherence
4.2: Decoherence
dephasing
4.2: Decoherence

E

entanglement
5.2: Entanglement
exchange integral
9.1: Interacting Electrons in Atomic Shells
expectation value
11.3.2: Uncertainty in Quantum Mechanics

F

ferromagnet
9.2 Spin Waves in Solids
Fock space
8.2: Creation and Annihilation Operators
free particle
13.3: Free Particles and the de Broglie Wavelength

H

Heisenberg picture
3: Schrödinger and Heisenberg Pictures
Higgs bosons
9.4: Outlook: Quantum Field Theory

I

indistinguishability
8.1: Symmetric and Anti-symmetric States
inner product
6.3: Bra Vectors and the Inner Product

K

Kraus' Representation Theorem
6.2: Positive and Completely Positive Maps

L

Ladder operators
8.2: Creation and Annihilation Operators
Lindblad equation
6.1: The Lindblad Equation

M

Many Worlds interpretation
2: The Postulates of Quantum Mechanics
measurement problem
2: The Postulates of Quantum Mechanics
mixed states
4.1: Mixed States

N

noise
4.3: Imperfect Measurements

O

orbital angular momentum
7.1: Orbital Angular Momentum
orthogonality
6.4: Normalization and Orthogonality

P

Pauli matrices
7.2: Spin
Pauli spin matrices
10.1: Spin Operators
Plane waves
13.3: Free Particles and the de Broglie Wavelength
Projection Operator
1.4: Projection Operators and Tensor Products

Q

quantum field theory
9.4: Outlook: Quantum Field Theory

R

rotating wave approximation
9.3: An Atom in a Cavity

S

Schrödinger picture
3: Schrödinger and Heisenberg Pictures
second quantization
8.2: Creation and Annihilation Operators
Slater determinant
8.1: Symmetric and Anti-symmetric States
spin
7.2: Spin
spin angular momentum
7.2: Spin
spin expectation values
10.2: Expectation Values
spin operators
10.1: Spin Operators
Spin waves
9.2 Spin Waves in Solids

T

tensor product
1.4: Projection Operators and Tensor Products
trace
1.5: The Trace and Determinant of an Operator

V

Von Neumann equation
4.1: Mixed States

W

wavefunction collapse
2: The Postulates of Quantum Mechanics

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Advanced Quantum Mechanics (Kok)

Webpages: 97

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC-SA 4.0: 96.9% (94 pages)
- Undeclared: 3.1% (3 pages)

By Page

- Advanced Quantum Mechanics (Kok) - *CC BY-NC-SA 4.0*
  - Front Matter - *CC BY-NC-SA 4.0*
    - TitlePage - *CC BY-NC-SA 4.0*
    - InfoPage - *CC BY-NC-SA 4.0*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Linear Vector Spaces and Hilbert Space - *CC BY-NC-SA 4.0*
    - 1.1: Linear Vector Spaces - *CC BY-NC-SA 4.0*
    - 1.2: Operators in Hilbert Space - *CC BY-NC-SA 4.0*
    - 1.3: Hermitian and Unitary Operators - *CC BY-NC-SA 4.0*
    - 1.4: Projection Operators and Tensor Products - *CC BY-NC-SA 4.0*
    - 1.5: The Trace and Determinant of an Operator - *CC BY-NC-SA 4.0*
  - 2: The Postulates of Quantum Mechanics - *CC BY-NC-SA 4.0*
  - 3: Schrödinger and Heisenberg Pictures - *CC BY-NC-SA 4.0*
  - 4: Mixed States and the Density Operator - *CC BY-NC-SA 4.0*
    - 4.1: Mixed States - *CC BY-NC-SA 4.0*
    - 4.2: Decoherence - *CC BY-NC-SA 4.0*
    - 4.3: Imperfect Measurements - *CC BY-NC-SA 4.0*
  - 5: Composite Systems and Entanglement - *CC BY-NC-SA 4.0*
    - 5.1: Composite Systems - *CC BY-NC-SA 4.0*
    - 5.2: Entanglement - *CC BY-NC-SA 4.0*
    - 5.3: Quantum teleportation - *CC BY-NC-SA 4.0*
  - 6: Evolution of Open Quantum Systems - *CC BY-NC-SA 4.0*
    - 6.1: The Lindblad Equation - *CC BY-NC-SA 4.0*
    - 6.2: Positive and Completely Positive Maps - *CC BY-NC-SA 4.0*
    - 6.3: Bra Vectors and the Inner Product - *CC BY-NC-SA 4.0*
    - 6.4: Normalization and Orthogonality - *CC BY-NC-SA 4.0*
    - 6.5: Interpreting the Inner Product - *CC BY-NC-SA 4.0*
      - 6.5.1: Propagating Amplitudes - *CC BY-NC-SA 4.0*
  - 7: Orbital Angular Momentum and Spin - *CC BY-NC-SA 4.0*
    - 7.1: Orbital Angular Momentum - *CC BY-NC-SA 4.0*
    - 7.2: Spin - *CC BY-NC-SA 4.0*
    - 7.3: Total Angular Momentum - *CC BY-NC-SA 4.0*
    - 7.4: Composite Systems with Angular Momentum - *CC BY-NC-SA 4.0*
  - 8: Identical Particles - *CC BY-NC-SA 4.0*
    - 8.1: Symmetric and Anti-symmetric States - *CC BY-NC-SA 4.0*
    - 8.2: Creation and Annihilation Operators - *CC BY-NC-SA 4.0*
    - 8.3: Observables Based on Creation and Annihilation Operators - *CC BY-NC-SA 4.0*
    - 8.4: Bose-Einstein and Fermi-Dirac Statistics - *CC BY-NC-SA 4.0*
  - 9: Many-Body Problems in Quantum Mechanics - *CC BY-NC-SA 4.0*
    - 9.1: Interacting Electrons in Atomic Shells - *CC BY-NC-SA 4.0*
    - 9.2 Spin Waves in Solids - *CC BY-NC-SA 4.0*
    - 9.3: An Atom in a Cavity - *CC BY-NC-SA 4.0*
    - 9.4: Outlook: Quantum Field Theory - *CC BY-NC-SA 4.0*
  - 10: Pauli Spin Matrices - *CC BY-NC-SA 4.0*
    - 10.1: Spin Operators - *CC BY-NC-SA 4.0*
    - 10.2: Expectation Values - *CC BY-NC-SA 4.0*
    - 10.3: Total Angular Momentum - *CC BY-NC-SA 4.0*

- 11: Noncommuting Operators and Uncertainty - *CC BY-NC-SA 4.0*
  - 11.1: Eigenstates and Commuting Operators - *CC BY-NC-SA 4.0*
  - 11.2: Non-Commuting Operators - *CC BY-NC-SA 4.0*
  - 11.3: Quantifying Uncertainty - *CC BY-NC-SA 4.0*
    - 11.3.1: Mean and Variance - *CC BY-NC-SA 4.0*
    - 11.3.2: Uncertainty in Quantum Mechanics - *CC BY-NC-SA 4.0*
  - 11.4: The Heisenberg Uncertainty Principle - *CC BY-NC-SA 4.0*
    - 11.4.1: An Alternate Formulation - *CC BY-NC-SA 4.0*
    - 11.4.2: Vacuum Energy - *CC BY-NC-SA 4.0*
- 12: Multiple Particle States - *CC BY-NC-SA 4.0*
  - 12.1: Indistinguishable Particles - *CC BY-NC-SA 4.0*
  - 12.2: Notating Multiple Particle States - *CC BY-NC-SA 4.0*
  - 12.3: The Exchange Operator - *CC BY-NC-SA 4.0*
  - 12.4: Fermions and Bosons - *CC BY-NC-SA 4.0*
  - 12.5: The Pauli Exclusion Principle - *CC BY-NC-SA 4.0*
  - 12.6: Entangled Particles - *CC BY-NC-SA 4.0*
- 13: The Schrödinger Equation - *CC BY-NC-SA 4.0*
  - 13.1: Where we are so far - *CC BY-NC-SA 4.0*
  - 13.2: Stating the Equation - *CC BY-NC-SA 4.0*
  - 13.3: Free Particles and the de Broglie Wavelength - *CC BY-NC-SA 4.0*
  - 13.4: Quantized Energy Levels in Bound Systems - *CC BY-NC-SA 4.0*
  - 13.5: The Simple Harmonic Oscillator - *CC BY-NC-SA 4.0*
  - 13.6: The Hydrogen Atom - *CC BY-NC-SA 4.0*
  - 13.7: Interpretation of the Wave Function ψ(x) - *CC BY-NC-SA 4.0*
- 14: Atomic Orbitals - *CC BY-NC-SA 4.0*
  - 14.1: The Schrödinger Equation - *CC BY-NC-SA 4.0*
  - 14.2: The Orbitals - *CC BY-NC-SA 4.0*
  - 14.3: Visualizing Orbitals - *CC BY-NC-SA 4.0*
    - 14.3.1: s Orbitals - *CC BY-NC-SA 4.0*
    - 14.3.2: p Orbitals - *CC BY-NC-SA 4.0*
    - 14.3.3: d Orbitals - *CC BY-NC-SA 4.0*
- 15: The Periodic Table of the Elements - *CC BY-NC-SA 4.0*
  - 15.1: Interacting Electrons, Energy Levels, and Filled Shells - *CC BY-NC-SA 4.0*
  - 15.2: Filling Up Orbitals - *CC BY-NC-SA 4.0*
  - 15.3: Reading a Periodic Table - *CC BY-NC-SA 4.0*
    - 15.3.1: Electronic Configuration - *CC BY-NC-SA 4.0*
- 16: Matter - *CC BY-NC-SA 4.0*
  - 16.1: The Standard Model of Particle Physics - *CC BY-NC-SA 4.0*
  - 16.2: Nuclei and Atoms - *CC BY-NC-SA 4.0*
  - 16.3: Molecules - *CC BY-NC-SA 4.0*
  - 16.4: Solids - *CC BY-NC-SA 4.0*
  - 16.5: Liquids, Gasses, and Plasmas - *CC BY-NC-SA 4.0*
    - 16.5.1: Quantum Gasses - *CC BY-NC-SA 4.0*
  - 16.6: Planets, Stars, Galaxies, and Clusters - *CC BY-NC-SA 4.0*
  - 16.7: Dark Matter and Dark Energy - *CC BY-NC-SA 4.0*
- Back Matter - *CC BY-NC-SA 4.0*
  - Index - *CC BY-NC-SA 4.0*
  - Glossary - *CC BY-NC-SA 4.0*
  - Detailed Licensing - *Undeclared*